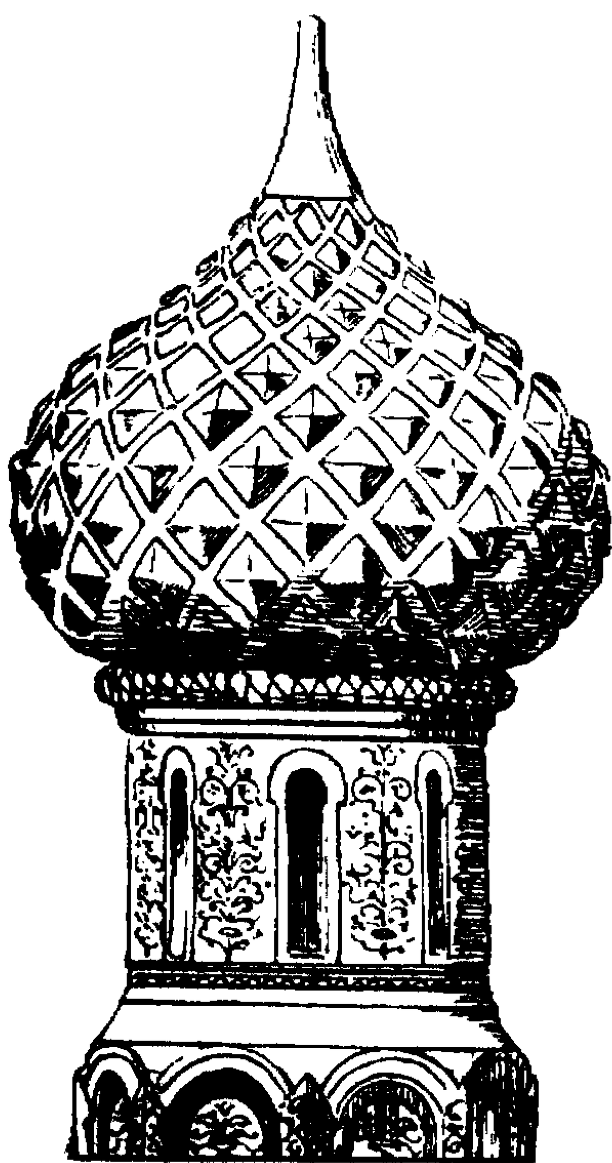


俄 罗 斯 数 学
教 材 选 译

● 数学天元基金资助项目

代数学引论 (第二卷)

线性代数 (第3版)

□ А. И. 柯斯特利金 著
□ 牛凤文 译


高等教育出版社
Higher Education Press

图字：01-2005-5733 号

А. И. Кострикин
Введение b алгебру. Часть II. Линейная алгебра, 2004

Originally published in Russian under the title
Introduction to Algebra
Part II: Linear Algebra by A. I. Kostrikin

图书在版编目（CIP）数据

代数学引论. 第2卷，线性代数：第3版/（俄罗斯）柯斯特利金著；牛凤文译. —北京：高等教育出版社，2008.1
（俄罗斯数学教材选译）
ISBN 978-7-04-021491-8

Ⅰ. 代…　Ⅱ. ①柯…②牛…　Ⅲ. ①代数-高等学校-教材 ②线性代数-高等学校-教材　Ⅳ. O15

中国版本图书馆 CIP 数据核字（2007）第 154118 号

策划编辑 赵天夫　**责任编辑** 蒋　青　**封面设计** 王凌波　**责任绘图** 郝　林
版式设计 王　莹　**责任校对** 姜国萍　**责任印制** 陈伟光

出版发行	高等教育出版社	**购书热线**	010—58581118
社　　址	北京市西城区德外大街 4 号	**免费咨询**	800—810—0598
邮政编码	100011	**网　　址**	http://www.hep.edu.cn
总　　机	010—58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京市鑫霸印务有限公司	**畅想教育**	http://www.widedu.com
开　　本	787×1 092　1/16	**版　　次**	2008 年 1 月第 1 版
印　　张	21.5	**印　　次**	2008 年 1 月第 1 次印刷
字　　数	430 000	**定　　价**	44.10 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

物料号　21491—00

《俄罗斯数学教材选译》序

从上世纪 50 年代初起, 在当时全面学习苏联的大背景下, 国内的高等学校大量采用了翻译过来的苏联数学教材. 这些教材体系严密, 论证严谨, 有效地帮助了青年学子打好扎实的数学基础, 培养了一大批优秀的数学人才. 到了 60 年代, 国内开始编纂出版的大学数学教材逐步代替了原先采用的苏联教材, 但还在很大程度上保留着苏联教材的影响, 同时, 一些苏联教材仍被广大教师和学生作为主要参考书或课外读物继续发挥着作用. 客观地说, 从解放初一直到文化大革命前夕, 苏联数学教材在培养我国高级专门人才中发挥了重要的作用, 起了不可忽略的影响, 是功不可没的.

改革开放以来, 通过接触并引进在体系及风格上各有特色的欧美数学教材, 大家眼界为之一新, 并得到了很大的启发和教益. 但在很长一段时间中, 尽管苏联的数学教学也在进行积极的探索与改革, 引进却基本中断, 更没有及时地进行跟踪, 能看懂俄文数学教材原著的人也越来越少, 事实上已造成了很大的隔膜, 不能不说是一个很大的缺憾.

事情终于出现了一个转折的契机. 今年初, 在由中国数学会、中国工业与应用数学学会及国家自然科学基金委员会数学天元基金联合组织的迎春茶话会上, 有数学家提出, 莫斯科大学为庆祝成立 250 周年计划推出一批优秀教材, 建议将其中的一些数学教材组织翻译出版. 这一建议在会上得到广泛支持, 并得到高等教育出版社的高度重视. 会后高等教育出版社和数学天元基金一起邀请熟悉俄罗斯数学教材情况的专家座谈讨论, 大家一致认为: 在当前着力引进俄罗斯的数学教材, 有助于扩大视野, 开拓思路, 对提高数学教学质量、促进数学教材改革均十分必要. 《俄罗斯数学教材选译》系列正是在这样的情况下, 经数学天元基金资助, 由高等教育出版社组织出版的.

经过认真选题并精心翻译校订, 本系列中所列入的教材, 以莫斯科大学的教材为主, 也包括俄罗斯其他一些著名大学的教材. 有大学基础课程的教材, 也有适合大学高年级学生及研究生使用的教学用书. 有些教材虽曾翻译出版, 但经多次修订重版, 面目已有较大变化, 至今仍广泛采用、深受欢迎, 反射出俄罗斯在出版经典教材方面所作的不懈努力, 对我们也是一个有益的借鉴. 这一教材系列的出版, 将中俄数学教学之间中断多年的链条重新连接起来, 对推动我国数学课程设置和教学内容的改革, 对提高数学素养、培养更多优秀的数学人才, 可望发挥积极的作用, 并起着深远的影响, 无疑值得庆贺, 特为之序.

李大潜

2005 年 10 月

“线性代数学既是一个古老的同时
又是一个全新的数学分支.”

H. 布尔巴基

序　　言

本书是整个《代数学引论》教程的第二卷(简记为[BA Ⅱ]), 它的目的在于系统地阐述数学的一个重要分支——线性代数学的基础, 尽管在本教程的第一卷中我们对其已有所触及. 因为代数理论的观点和几何理论的观点同等重要, 因此, 线性代数学和几何学这一对典型的“孪生姐妹”将会以同样的身份呈现出来. 在平面和三维空间的解析几何教程中已经知道了很多对于两个或者三个变元的代数关系式的几何解释. 重要的是, 线性代数依据几何直观支撑的术语和概念适用于任意维数n的n维空间.

“线性代数与分析”, “线性代数与微分方程”以及其他更多在大学教程中使用的术语反映出这样一个事实, 线性的概念是数学中最为普及的概念之一, 或者, 更广泛地说, 它是整个自然科学中最基本的概念之一. 把问题分成线性的和非线性的并不是要满足数学家们的特殊癖好, 而是在更广泛意义上理解的线性代数力所不及的地方, 我们的直观的相对弱点所造成的, 这一点我们已经完全认识清楚了.

在20世纪初就已经完全发育成型的线性代数体系在不同的方向上继续得到发展且日臻完美. 与此同时, 它的依赖于极限过程的无穷维部分, 本质上说, 走向了泛函分析, 而计算部分, 特别是与实际使用电子计算机的可能性相关的部分, 变成了独立的科学的研究对象. 现在提供的这本书不可能充当面面俱到的线性代数手册, 这不仅仅是因为它不能包括上面提到的两个方向, 而首先是因为它对应用的阐述不够充分(尽管这最后一章可以称为是应用). 在这方面, 列在补充文献清单中的[2]包含了

丰富得多的想象力, 令人沉思的动议, 并且还有线性代数概念的量子力学解释. 可以把该书推荐给所有想要向标准教程范围以外涉猎的读者. 而现在这本书最多只能容纳[2]中的一些片断. 我们的打算和希望可以归结为, 读者(首先是一年级大学生)详细地研读了教科书内容(在一个学期内每周四个小时上课, 四个小时练习), 其后再以两书的补充章节为家庭读物, 能在线性代数领域培养出现代数学思想.

不言而喻, 为了理解本教科书的课文只需要很好地掌握第一卷(简记为[BA I]), 也就是第一个学期的学习内容. 这两卷中的术语和表达是完全一致的, 而所有新引进的东西都做了特别的解释, 顺便说一句, 第p章§q的习题r在课文中总是简单地表达成习题p. q. r. 不同于[BA I, BA III], 用专门篇幅单列的习题提示与答案读者最好到万不得已时再去光顾它.

作者清楚地认识到, 把教学参考书[2]"变得通俗化", 特别是强行改变其固有的格调是一次极其徒劳无益的任务. 现在(却)这样阐述了, 仅有的可辩解的理由是作为大学生质疑的回应, 这种适应早就准备好了, 而仅仅由外部原因大大延迟未能实现罢了.

补充文献

1. *Кострикин А.И.* Введение в алгебру. Ч.I. Основы алгебры. —2-е изд.—М.: Наука, 1994.—318 с.

即本教科之第一卷, 张英伯译, 高等教育出版社出版.

2. *Кострикин А.И., Манин Ю.И.* Линейная алгебра и геометрия.—М.: Наука, 1986. —304 с.

3. Сборник задач по алгебре/Под ред. А.И.Кострикина. —М.: Факториал, 1995.—456 с.

4. *Гельфанд И.М.*Лекции по линейной алгебре. —5-е изд.—М.: Наука, 1998. —272 с.

中译本: И.М. 盖尔冯德. 一次代数学. 第二版(1950). 刘亦珩译. 商务印书馆, 1953年.

И. М. 盖尔冯德. 线性代数学, 刘亦珩译. 高等教育出版社. 1957年.

5. *Мальчев А.И.*Основы линейной алгебры. —М.: Наука, 1956. —340 с.

中译本: А.И. 马力茨夫. 线性代数基础. 第二版(1956). 柯召译. 高等教育出版社, 1959年.

6. *Халмош П.Р.* Конечномерные векторные пространства.—М.: Мир, 1970.—264 с.

7. *Артин Э.* Геометрическая алгебра. —М.: Мир, 1970. —284 с.

8. *Шилов Г.Е.* Введение в теорию линейных пространств. —М.: Наука, 1956.—304 с.

中译本: Г.Е.希洛夫. 线性空间引论, 周学光等译. 高等教育出版社, 1957.

9. *Фаддеев Д.К., Фаддеева В.Н.* В ычислительные методы линейной алгебры. —М.: Наука, 1963.

10. *Стренг Г.*Линейная алгебра и её применения.—М.:Мир,1980.—454 с.

11. *Прасолов В.В.* Задачи и теоремы линейной алгебры. —М.: Наука, 1991.

12. *Беллман Р.* Введение в теорию матриц. —М.: Наука, 1976. —368 с.

13. *Гантмахер Ф.Р.* Теория матриц. —М.: Наука, 1967.

中译本: Ф.Р. 甘特马尔赫. 矩阵论. 1953年版. 上下卷. 柯召译. 高等教育出版社, 1955年.

14. *Ланкастер П.* Теория матриц. —М.: Наука, 1978.

15. *Huppert B.* Angewandte Lineare Algebra. —Berlin-New York: Walter de Gruyter, 1990.—646 p.

目　录

第1章　空间与形式

也许未必值得指出，没有经过开发的森林不同于有人侍弄的公园或者人工栽培的井然有序的树林. 在所有这些差别中又含有那么多的相同之处，以至于在不能品尝蘑菇味道，不能欣赏修剪过的草坪的魅力的外星人看来，林区就是连成一片的，长满草木的，充满各种不同高度和形状的山体，而我们却称之为大森林. 如果把本书的这一章与[BA Ⅰ]中讲述坐标向量空间的第2章相比较，就会发生类似的事情. 抽象线性空间是用公理化方法引进的，它的元素被称为向量，正因如此，也经常称它为向量空间. 相应的公理系统，本质上仍然是G. 佩亚诺(1888)年完成的，它很好地适应了在线性代数中占有中心地位的线性映射(特别地，线性算子)理论的需要. 与此同时，矩阵的概念似乎也不再居次要地位了，首要的意义在于获得了研究对象不依赖于基底选择的不变性质.

但是，在深入到抽象的森林去之前，不妨再一次沿着有人侍弄的公园走一走，即回顾一下由具体的长度为n的行向量组成的空间. 我们有意地在已知资料的部分的重复中进行，以磨平抽象叙述的不顺畅之处.

§1　抽象向量空间

1. 论据与公理系统　在[BA Ⅰ]的第2章，我们研究了长度为n的行构成的n维向量空间$\mathbb{R}^n = \{(x_1, \cdots, x_n) | x_i \in \mathbb{R}\}$以及与$m \times n$矩阵互为单值对应的$\mathbb{R}^n \to \mathbb{R}^m$线性映射，当$m = n$时，双射线性映射$\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$的特点是行列式性质$\det A \neq 0$，后者允许用克拉默规则求解把$\varphi_A$与$\mathbb{R}^n$中一个固定向量结合在一起的那个线性方程组，在$\det A \neq 0$的情形，齐次线性方程组构成$\mathbb{R}^n$的一个子空间，但是，正如当时就已经指出的那样，这个子空间(更精确些，线性包络)是另有规律的对象：如果允许$\mathbb{R}^n$有基底$(1, 0, \cdots, 0), \cdots, (0, \cdots, 0, 1)$，那么线性包络$U \subset \mathbb{R}^n$通常并不具有这样形式的

基底, 这种不方便是由于$\mathbb{R}^n$的规律太过于具体造成的.

事实上, 我们已经实实在在地使用过的, 并将在下面反复出现的性质$\text{B}\Pi_1$—$\text{B}\Pi_8$不只适用于空间$\mathbb{R}^n$, 例如, 看在中学已经研究过的微分方程$d^2x/dt^2+x=0$. 显然, 它的一般解可写成$x(t)=\alpha\sin t+\beta\cos t$的形式. 如果$\alpha_0,\beta_0$使得对所有的$t$都有$\alpha_0\sin t+\beta_0\cos t=0$, 那么, 令$t_1=\pi/2$, $t_2=0$就得到$\alpha_0=0=\beta_0$. 这个情况为我们做下面的事情提供了基础: 用以下定义1的精神去讨论部分解$\sin t$, $\cos t$的线性相关性和讨论方程式$d^2x/dt^2+dx=0$一般解的二维线性空间.

定义1 设$\mathfrak{K}$是个任意域. 满足下列公理的元素(叫做**向量**)的集合V称之为域$\mathfrak{K}$上的**向量**(或**线性**)**空间**:

i) 在V上给定了一个二元运算$V\times V\to V$, 通常记成加: $(\mathbf{x},\mathbf{y})\mapsto\mathbf{x}+\mathbf{y}$, 且赋予$V$阿贝尔群结构(**空间$V$的加法群**). 可见

$\text{B}\Pi_1$: $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$(**交换律**);

$\text{B}\Pi_2$: $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$(**结合律**);

$\text{B}\Pi_3$: 存在一个与众不同的元素$\mathbf{0}$, 称为**零向量**, 它对任意$\mathbf{x}\in V$都有$\mathbf{x}+\mathbf{0}=\mathbf{x}$;

$\text{B}\Pi_4$: 对每个$\mathbf{x}\in V$存在一个**逆**(或**反**)向量$-\mathbf{x}$使得$\mathbf{x}+(-\mathbf{x})=\mathbf{0}$.

ii) 在集合$\mathfrak{K}\times V$上给定一运算$(\lambda,\mathbf{x})\mapsto\lambda\mathbf{x}$, 称为用$\mathfrak{K}$中的纯量乘$V$中的向量而且要具有下列性质.

$\text{B}\Pi_5$: $1\cdot\mathbf{x}=\mathbf{x}$(**酉性**);

$\text{B}\Pi_6$: $(\alpha\beta)\mathbf{x}=\alpha(\beta\mathbf{x})$对任意$\alpha$, $\beta\in\mathfrak{K}$和$\mathbf{x}\in V$都成立(**结合律**);

加法和乘法用下面的两个分配律联系在一起:

$\text{B}\Pi_7$: $(\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$;

$\text{B}\Pi_8$: $\lambda(\mathbf{x}+\mathbf{y})=\lambda\mathbf{x}+\lambda\mathbf{y}$.

要注意这样一种情况, 在$\text{B}\Pi_7$的左侧的符合+是对域$\mathfrak{K}$的元素而言的, 而右侧的符号+则是对于向量的. 更严格地说, 应该用不同的符号表示在阿贝尔群V中的加法和域$\mathfrak{K}$中的加法(比如说, 用$\oplus$和+). 同样的处理还有$\mathfrak{K}\times V$上的乘法运算和域$\mathfrak{K}$的乘法运算(比方说, 用$\odot$和$\cdot$). 但是, 通常并不这样做, 因为它们各自的作用总是非常清楚的, 然而, 为了使这个说明内容更丰富且预先避免出错, 我们研究一下正实数的集合$V=\mathbb{R}_+$. 令$x\oplus y=xy$ (集合$\mathbb{R}$中的通常的乘法)和$\lambda\odot x=x^\lambda$(x的λ次幂, $x\in\mathbb{R}_+,\lambda\in\mathbb{R}$)就不难相信, 公理$\text{B}\Pi_1-\text{B}\Pi_8$是成立的, 所以$V$是$\mathbb{R}$上的向量空间. $1\in\mathbb{R}_+$是零向量. 显然, 在给出的这种情形通常的记法$x+y=xy,\lambda x=x^\lambda$就有可能引起困惑.

这里还有一个例子, 另外一种更令人喜欢的方法. 设V是复数$\mathbb{C}$上的一个向量空间, 我们定义一个新的向量空间$\bar{V}$, 它与V具有相同的加法子群V, 但有另外一种用纯量做的乘法$(\lambda,\mathbf{x})\mapsto\lambda\odot\mathbf{x}=\bar{\lambda}\mathbf{x}$, 其中$\bar{\lambda}$是$\lambda$的共轭复数, 因为$\lambda\mapsto\bar{\lambda}$是域$\mathbb{C}$上的自同构, 容易相信$\bar{V}$是个向量空间. 不用符号$\odot$(或者其他的某个符号)同时研究$V$和$\bar{V}$, 可能还是挺困难的.

约定 读者可能已经注意到了，在我们这里，空间V的向量用黑体，小写拉丁字母表示，而有时也用希腊字符. 但是，在抽象向量空间中将用普通斜体字; 在具体的特例中遵循这一原则可能是不便于实践的，而用符号头顶上加箭头表示向量又太笨重了. 在并不复杂的技巧之下，我们采用的妥协就不会引起误会.

由向量空间V的定义可以直接导出一些推论，在今后使用它们时就不再引证了:

i) 对所有的$\lambda \in \mathfrak{K}, \mathbf{x} \in V$都有$0\mathbf{x} = \lambda\mathbf{0} = \mathbf{0}$. 事实上，由$\text{B}\Pi_7$, $0\mathbf{x} = (0+0)\mathbf{x} = 0\mathbf{x} + 0\mathbf{x}$, 从而$0\mathbf{x}=\mathbf{0}$. 类似的，$\lambda\mathbf{0} = \lambda(\mathbf{0}+\mathbf{0}) = \lambda\mathbf{0} + \lambda\mathbf{0}$, 也就是$\lambda\mathbf{0} = \mathbf{0}$;

ii) $\lambda\mathbf{x} = \mathbf{0} \Rightarrow \lambda = 0$或者$\mathbf{x}=\mathbf{0}$. 如果，比方说，$\lambda \neq 0$, 那么$\mathbf{x}=1\cdot\mathbf{x}=(\lambda^{-1}\lambda)\mathbf{x}=\lambda^{-1}(\lambda\mathbf{x})=\lambda^{-1}\mathbf{0}=\mathbf{0}$;

iii) $(n\cdot 1)\mathbf{x} = \mathbf{x} + \mathbf{x} + \cdots + \mathbf{x}$($n$重加). 简单地用$n\mathbf{x}$代替$(n\cdot 1)\mathbf{x}$是很自然的，其中1是域$\mathfrak{K}$的单位. 如果$\mathfrak{K}$是个特征数为$p$的有限域，那么$p\mathbf{x}=\mathbf{0}$;

iv) $(-1)\mathbf{x} = -\mathbf{x}$. 事实上，$\mathbf{x} + (-1)\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = (1+(-1))\mathbf{x} = 0\mathbf{x} = \mathbf{0}$

2. 线性包络. 子空间 注意，给定任意一整套纯量$\lambda_1, \cdots, \lambda_n \in \mathfrak{K}$和向量$\mathbf{x}_1, \cdots, \mathbf{x}_n$, 我们可以建立表达式

$$\lambda_1\mathbf{x}_1 + \cdots + \lambda_n\mathbf{x}_n = \sum_{i=1}^{n} \lambda_i\mathbf{x}_i$$

并称其为向量$\mathbf{x}_i$以λ_i为系数的**线性组合**. 更一般地，如果I是某个指标集，它可能是无限的，且$M = \{\mathbf{x}_i \in V | i \in I\}$是$V$中向量的子集，那么，对任意系数$\lambda_i \in \mathfrak{K}$, 而它们之中只有有限个不为零，研究其线性组合$\sum_{i\in I} \lambda_i\mathbf{x}_i$都是有意义的，显然

$$\lambda\left(\sum \lambda_i\mathbf{x}_i\right) = \sum (\lambda\lambda_i)\,\mathbf{x}_i$$

是以$\lambda\lambda_i$, $i \in I$, 为系数的线性组合. 类似地，两个分别以λ_i和μ_i为系数的同样的向量$\mathbf{x}_i$的线性组合的和

$$\left(\sum_{i\in I} \lambda_i\mathbf{x}_i\right) + \left(\sum_{i\in I} \mu_i\mathbf{x}_i\right) = \sum_{i\in I} (\lambda_i + \mu_i)\mathbf{x}_i$$

就是向量$\mathbf{x}_i \in M$的以$\lambda_i + \mu_i$为系数的线性组合. 这些$\lambda_i+\mu_i$中只有有限个不为零. 这样一来，$\mathbf{x}_i \in M$的各种可能的线性组合的集合$\langle M\rangle_{\mathfrak{K}}$相对于向量加法运算和向量乘以纯量的运算是封闭的:

$$\lambda \in \mathfrak{K}, \mathbf{x}, \mathbf{y} \in \langle M\rangle \Rightarrow \mathbf{x} + \mathbf{y} \in \langle M\rangle, \lambda\mathbf{x} \in \langle M\rangle.$$

简法说$\langle M\rangle$是集合$M \subset V$的**线性包络**.

定义2 设V是域$\mathfrak{K}$上的向量空间，$U \subset V$是它的一个子集，而且它还是V的加法子群，同时，在纯量乘之下回到自身，那么，V上运算在U上的限制导出U的向量空间结构，把它称为V的**向量**(或**线性**)**子空间**.

任意多个向量子空间的交集仍然是向量子空间(见§2第4目的开头; 这是个简单的练习, 对于群的情形的练习在[BA Ⅰ]中已经见过了). 我们看到, 向量族$M \subset V$的线性包络是V的子空间. 而且, 显然, $\langle M\rangle$是包含M的最小的子空间. 还可以说, $\langle M\rangle$是由$\mathbf{x} \in M$**张成**的子空间或者是由$\mathbf{x} \in M$**生成**的子空间. 如果, 一开始M本身就是一个子空间, 那么, $\langle M\rangle = M$.

我们引入一些今后要遇到的向量空间的例子.

例1(零维空间) 在任意域$\mathfrak{K}$上都存在一个零维(一个元素)向量空间$V = \{\mathbf{0}\}$, 它的用纯量做乘法的规律是$\lambda\mathbf{0} = \mathbf{0}$.

例2(将基础域本身作为一维坐标空间) 定义$V = \mathfrak{K}$, V的基本运算与域$\mathfrak{K}$的运算一致. 如果1是域的单位元, 那么, 可以认为, $\mathfrak{K} = \langle 1\rangle$是1张成的线性包络.

更一般地: 如果域$\mathfrak{K}$是自己的一个子域$\mathfrak{B}$的扩域, 那么$\mathfrak{K}$就可以看成是$\mathfrak{B}$上的向量空间. 例如, 复数域$\mathbb{C}$是实数域$\mathbb{R}$上的向量空间, 而$\mathbb{R}$是有理数域$\mathbb{Q}$上的向量空间.

例3(n维坐标空间$\mathfrak{K}^n$; 见[BA Ⅰ]第2章, 其中$\mathbb{R}$可用任意域$\mathfrak{K}$替换) 当n=1时, 就得到上面的例子. 我们很快就可以看到(见§3), 所有的子空间$U \subset \mathfrak{K}^n$必然是某个齐次线性方程组的解构成的空间.

例4(函数空间) 在[BA Ⅰ] 第4章§3之第1目已经引进了函数环K^X, 它实际上还是K上线性空间(环K需要用域来代替). 换言之, X是任意一个集合, $\mathfrak{K}$是个域, $\mathfrak{K}^X$是所有映射(函数): $f : X \to \mathfrak{K}$的集合, 它具有逐点定义的加法和用纯量乘:

$(f + g)(x) = f(x) + g(x)$, 对所有的$x \in X$;

$(\lambda f)(x) = \lambda(f(x))$, 对所有$\lambda \in \mathfrak{K}, x \in X$.

可以对每个$x \in X$建立一个相应的称之为集中在$\{x\}$上的Δ函数δ_x:

$$\delta_x(x) = 1, \delta_x(x') = 0, x \neq x'.$$

如果$X = \{1, 2, \cdots, n\}$, 那么, 通常用δ_{ij}取代$\delta_i(j)$, 这是对于克罗内克符号的标准表达. 在这种情形, $\mathfrak{K}^X$与$\mathfrak{K}^n$相同. 换言之, 函数f可以相应地表成它的所有的值的向量$(f(1), f(2), \cdots, f(n))$, 而函数本身可单值地表成$\Delta$函数的线性组合形式

$$f = f(1)\delta_1 + f(2)\delta_2 + \cdots + f(n)\delta_n.$$

对于无穷集合X的情形, 类似的结论失去了意义, 因为无穷多个向量的和没有定义(如果没有特别涉及拓扑学的话).

可以大致相似地研究定义在整个直线或者区间$(a, b) \subset \mathbb{R}$上的实值函数. 容易验证, 所有连续函数的空间$\mathbb{R}^{(a,b)}_{\text{cont}}$作为一个子空间包含在线性空间$\mathbb{R}^{(a,b)}$中, 所有的连续可微的函数的空间$\mathbb{R}^{(a,b)}_{\text{diff}}$也作为子空间包含在$\mathbb{R}^{(a,b)}$中, 等等, 因此, 所有已经指出过的性质在函数作加法和用纯量作乘时都会继续保持.

例5 次数$\leqslant n - 1$的多项式$f \in \mathfrak{K}[t]$对通常的多项式加法和用纯量乘多项式的运算构成一个向量空间P_n. 应该指出, 次数等于一个固定数k的所有多项式不能构成一个线性空间. 但m个变元的k次型与零一起加以研究, 构成一个向量空间.

例6 设$g(t)$是一个固定的在区间[0, 1]上连续的实函数, 在某个区间$J \subset [0,1]$上不等于零. 而$V_n(g)$是所有形如$f(t)g(t)$的函数的集合, 其中$f(t)$是次数$\leqslant n-1$的多项式, 那么$V_n(g)$是个包含在$\mathbb{R}^J_{\text{cont}}$中的向量空间.

例7(矩阵空间) 按矩阵计算规则(见[BA I]第2章), 任意阶数为$m \times n$的矩阵可用域$\mathfrak{K}$中的元素乘之, 且任意两个相加结果仍得到同样类型的矩阵, 且满足这里所有的公理, 所以$m\times n$阶矩阵构成一个向量空间. 当$m = n$时, 方阵环$M_n(\mathfrak{K})$同时也是$\mathfrak{K}$上的向量空间, 被称为**代数**. 相应的一般定义我们将在第2章§2给出, 而且按实际作用, 在例2和例4的对象中经常碰到.

例8[Amer. Math. Monthly, 1990, V. 94, p60—p62] 矩阵$A \in M_n(\mathbb{Q})$被称为是**半幻的**(或者是**半幻方**), 如果矩阵的每一行的系数的和与每一列的系数的和都相同:

$$\sum_{k=1}^{n} a_{ik} = \sum_{k=1}^{n} a_{kj} = \sigma(A) \in \mathbb{Q}, 1 \leqslant i, j \leqslant n.$$

如果还有$\operatorname{tr} A = \sigma(A) := \sum\limits_{i=1}^{n} a_{i,n+1-i}$,那么, 就称$A$是**魔幻的**(或者**幻方**).

幻方自古以来就吸引人们的注意力. 它引起我们关注的一个相当明显的原因是, 所有半幻方的集合$\mathrm{SMag}_n(\mathbb{Q})$, 以及同样的, 所有幻方的集合$\mathrm{Mag}_n(\mathbb{Q})$都是$\mathbb{Q}$上的向量空间, 而且

$$\mathrm{Mag}_n(\mathbb{Q}) \subset \mathrm{SMag}_n(\mathbb{Q}) \subset M_n(\mathbb{Q}).$$

值得一说的是, 我们没有研究以正整数$1, 2, \cdots, n^2$为系数的$n \times n$幻方组合集合的数目. 关于这点, 见M. M. 鲍斯基尼柯夫《幻方》(Магические квадраты), M.: Hayka, 1964.

3. 关于几何解释的说明 如果$\mathfrak{K} = \mathbb{R}$(对应地, $\mathfrak{K} = \mathbb{C}$), 我们就称向量空间$V$是**实的**(对应地, **复的**). 也就是说, 从实践的观点看, 这种情形最令人感兴趣, 尽管大量的理论部分并不取决于域$\mathfrak{K}$的特性.

在我们生存的三维空间中, 由固定点引出的所有线段的集合, 毫无疑问, 可以充当向量空间的极自然的模型. 用数$\lambda \in \mathbb{R}$来乘一个线段, 即使得它的长度延长$\lambda > 1$倍(或$\lambda \leqslant 1$, 压缩λ倍), 并且λ为负数时取反方向. 有向线段的加法按平行四边形法则建立. 这个实向量空间同样与自由几何向量的集合一致, 只要约定, 两个有向向量经平移后能重合即为相等.

物理的3维空间$\mathbb{R}^3_\Phi$的对象可以借助图形加以表达. 多维空间的情形(我们将在§2讨论维数), 我们直觉能力受到严峻考验, 然而, 系统地对几何方式的呼应不仅是有用的而且是必要的: 形成稳定成型的结合, 使得理论更加生机勃勃.

对向量空间几何学感到生疏可能也与域$\mathfrak{K}$的特性有关系. 例如, $\mathfrak{K} = \mathbb{C}$, 那么在$\mathbb{C}$上的直线就是1维坐标空间$\mathbb{C}^1$, 复平面$\mathbb{R}^2$(不要将其与$\mathbb{C}^2$混淆)可作为它的直观的几何表示. 数$z = x + iy \in \mathbb{C}^1$对应点$(x, y) \in \mathbb{R}^2$, 用数$a \neq 0$乘之, 得到$|a|$倍延长并逆

时针旋转角度arg a. 特别地, 当$a=-1$时, 它在$\mathbb{R}^1$上的限制把$\mathbb{C}^1$旋转180° 给出"翻转"直线. 在第3章§4将叙述复化和实化运算, 在其中利用了代数封闭域$\mathbb{C}$的优越性, 而n维复空间$\mathbb{C}^n$就可以表示成$2n$维的实空间$\mathbb{R}^{2n}$.

还要注意, 物理空间$\mathbb{R}^3_\Phi$比同样维数的坐标空间内容丰富得多, 因为在$\mathbb{R}^3_\Phi$中定义了向量的长度, 向量间的角度, 图形的面积和体积. 所有这些补充信息不知不觉地转移到用以反映抽象空间的图形上, 这个抽象空间的公理系统暂时还是贫乏的, 以度量概念丰富起来的公理系统仅在后面各章才完全地实现.

至于向量空间V在何种程度上受纯量域性质的影响也同样可以从以下看出: 如果$\mathfrak{K}$是个有限域, 那么由$\mathfrak{K}^3$带来的几何结构是"有窟窿的"($\mathfrak{K}$的离散性的推论). 但$\mathfrak{K}$ 的这个缺点有时可以结合$\mathfrak{K}$上的线性几何学使用另一种离散图(见图1). 例如, 在二元域$\mathbb{F}_2=\{0,1\}\cong\mathbb{Z}_2$上的$n$维坐标空间$\mathbb{F}_2^n$允许与$\mathbb{R}^n$中的$n$维立方体顶点的集合$\{(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n);\varepsilon_i=0$或者$\varepsilon_i=1\}$自然地等同起来.

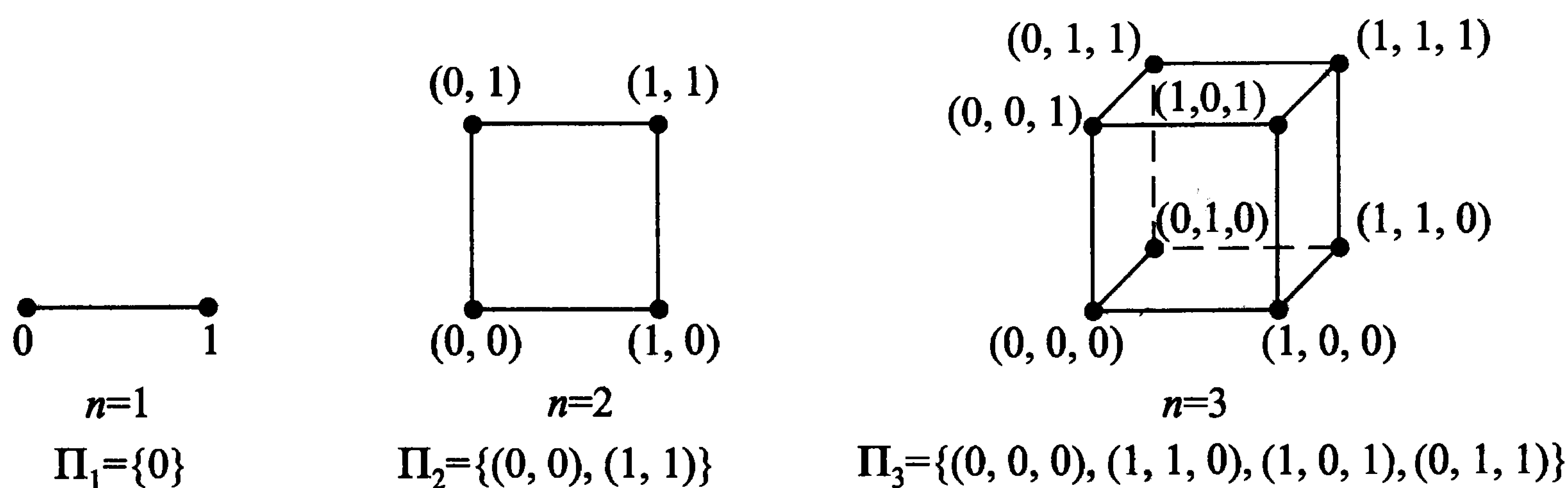


图1

由具有性质$\varepsilon_1+\varepsilon_2+\cdots+\varepsilon_n=0$的点$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$构成的子空间$\Pi_n$(注意1+0=0+1=1; 0+0=0=1+1)给出一个最简单的可以校正一个错误的编码(见[BA Ⅰ], 第4章, §3之第6目). 也就是说, 约定, 编码信号只对应点$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\in\Pi_n$, 同时, 取得具有性质$\varepsilon'_1+\varepsilon'_2+\cdots+\varepsilon'_n=1$的点$(\varepsilon'_1,\varepsilon'_2,\cdots,\varepsilon'_n)$, 我们有充分的根据可以认为, 在接收时, 信息带着外部干扰所导致的错误一起过来. 我们的码带有奇偶性验证, 自然地, 不能发现两个错误, 因为, 那时$(\varepsilon'_1,\varepsilon'_2,\cdots,\varepsilon'_n)\in\Pi_n$.

习 题

1. 下列集合是否构成域$\mathbb{R}$上的向量空间:

(1) $M_n(\mathbb{R})$中秩为一固定数r的所有矩阵;

(2) $M_n(\mathbb{R})$中所有对称矩阵(${}^tA=A$);

(3) $M_n(\mathbb{R})$中所有斜对称矩阵(${}^tA=-A$);

(4) $M_n(\mathbb{R})$中所有行列式为零的矩阵;

(5) $M_n(\mathbb{R})$中所有迹为零$\operatorname{tr}A=0$的矩阵(矩阵$A=(a_{ij})$的迹用关系式$\operatorname{tr}A=a_{11}+a_{22}+\cdots+a_{nn}$定义);

(6) $M_n(\mathbb{R})$中所有迹为正数的矩阵;

(7) 对固定的n, 所有形如$f(t) = (a_1 + a_2 + \cdots + a_n) - a_1 t - \cdots - a_n t^n$的多项式($a_i \in \mathbb{R}$).

2. 在有限的p元域$\mathbb{F}_p$上的长度为n的行向量$(x_1, \cdots, x_n)$的坐标空间$\mathbb{F}_p^n$中共有多少个元素? 在$\mathbb{F}_p^n$中方程$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$($\alpha_i \in \mathbb{F}_p$不全等于零)有多少个解?

§2 维数与基底

1. 线性相关性 类似于[BA Ⅰ, 第2章], 我们引进下面的

定义1 空间V的向量$\mathbf{v}_1, \cdots, \mathbf{v}_n$称之为**线性相关**的, 如果它们有一个非平凡的线性组合存在; 换言之, 有不全为零的纯量$\alpha_1, \cdots, \alpha_n$使得

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n = \mathbf{0}.$$

否则, 就称向量组$\mathbf{v}_1, \cdots, \mathbf{v}_n$是**线性无关**的.

当向量$\mathbf{v}_1, \cdots, \mathbf{v}_n$中有一个是零, 这些向量不可能是线性无关的: 例如, $\mathbf{v}_1 = \mathbf{0}$, 那么, 令$\alpha_1 = 1, \alpha_2 = \cdots = \alpha_n = 0$, 即得到一个等于零的非平凡线性组合.

定理1 向量$\mathbf{v}_1, \cdots, \mathbf{v}_n, n \geqslant 2$是线性相关的, 当且仅当, 其中之一个向量是其余向量的线性组合. 向量组线性相关, 只要它的某一个部分组是线性相关的, 换言之, 如果向量组线性无关, 那么, 它的所有的部分组都是线性无关的.

我们有意省略了证明, 因为它是足够显然的而且是[BA Ⅰ]第2章§1中相应定理证明的重复, 事实上, 随后可以证明

定理2 如果在空间V中, 线性无关的向量组$\mathbf{e}_1, \cdots, \mathbf{e}_s$的每个向量都是向量$\mathbf{f}_1, \cdots, \mathbf{f}_t$的线性组合, 那么, $s \leqslant t$.

证明 回想我们以前的推断(仅仅相差一点表达). 按条件有

$$\mathbf{e}_1 = \alpha_{11}\mathbf{f}_1 + \alpha_{21}\mathbf{f}_2 + \cdots + \alpha_{t1}\mathbf{f}_t,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\mathbf{e}_s = \alpha_{1s}\mathbf{f}_1 + \alpha_{2s}\mathbf{f}_2 + \cdots + \alpha_{ts}\mathbf{f}_t,$$

其中α_{ij}是纯量. 设$s > t$. 用系数x_j作成$\mathbf{e}_1, \cdots, \mathbf{e}_s$的线性组合

$$\begin{aligned} & x_1\mathbf{e}_1 + \cdots + x_s\mathbf{e}_s \\ = & (\alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1s}x_s)\mathbf{f}_1 + \cdots + (\alpha_{t1}x_1 + \alpha_{t2}x_2 + \cdots + \alpha_{ts}x_s)\mathbf{f}_t. \end{aligned}$$

进一步, 看由具有s个未知量的t个方程组成的线性方程组

$$\alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1s}x_s = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_{t1}x_1 + \alpha_{t2}x_2 + \cdots + \alpha_{ts}x_s = 0.$$

那么, 按假定, 我们的齐次方程组必有非零解$(\beta_1, \cdots, \beta_s)$(参阅相关的[BA Ⅰ] 第1章§3推论2及[BA Ⅰ]第4章§3, 在那里得到了任意纯量域上线性系统的重要见解). 这意味

着, $\beta_1\mathbf{e}_1+\cdots+\beta_s\mathbf{e}_s=\mathbf{0}$是个非平凡的线性关系式, 但这与定理的条件相矛盾, 所以$s\leqslant t$. □

推论 *V中任意两个等价的线性无关组都必有同样数量的向量(可以是无限多个).*

这里, 我们认为两个向量组是**等价**的, 如果一个组里的每一个向量都是另一向量组的线性组合.

当然, 等价的线性相关的向量组可以分别由不同数量的向量组成, 甚至, 一个等价组可以是线性无关的, 而另一个是线性相关的, 然而, 如果在V中的给定的向量组中我们取两个极大的线性无关的部分组(**极大**意味着不可能由更多的向量将其扩充成线性无关的部分组), 那么, 这两组必然有同样多个向量. 可用定理1和定理2证明之.

定义2 *在给定的向量组中任意一个极大线性无关的部分组所含的向量的个数被称为该向量组的**秩**(或称**秩数**).*

我们建立的应用于空间V上的事实还有另外一些解释, 这些解释将在今后的叙述中起奠基性作用.

2. 向量空间的维数与它的基底 可以分成两种情况: 或者在空间V中有任意多个向量线性无关(有任意秩的向量组), 此时, 说V是**无穷维**的; 或者V中所有足够多的向量作成的向量组都线性相关. 无穷维线性空间的内容丰富的理论要求具备更多的附加条件, 通常是拓扑结构, 这里只会偶尔地遇到无穷维空间.

定义3 *在线性空间V中有n个线性无关的向量, 而不存在具有更大数目向量的线性无关系, 则称V是n**维**(记为$\dim_{\mathfrak{K}}V=n$或者更简单地$\dim V=n$). 零空间被认为是**零维**的.*

这个定义与直线(1维空间), 平面(2维空间)和空间$\mathbb{R}^3_\Phi$(3维空间)的维数概念是完全一致的.

用新的术语, 向量组$\{\mathbf{v}_1,\mathbf{v}_2,\cdots\}$的秩与线性包络$\langle\mathbf{v}_1,\mathbf{v}_2,\cdots\rangle$的秩相同.

例 i) 坐标空间$\mathfrak{K}^n$维数为n(见[BA Ⅰ]第2章), 要不是这样, 我们这里定义的维数概念就是不完备的.

ii) $m\times n$阶矩阵空间的维数是mn, 在这个空间中, 把矩阵的元素排成长度为mn的一行, 就可以用坐标空间$\mathfrak{K}^{mn}$来标记$m\times n$阶的矩阵空间.

iii) §1中例4中的函数空间显然是无穷维的.

iv) 一个变元的次数$\leqslant n-1$的多项式空间P_n显然是n维的, 比如, 向量$1,t,\cdots,t^{n-1}$就是线性无关的.

v) m个变元的k次齐次型空间的维数是$n=\begin{pmatrix}k+m-1\\k\end{pmatrix}$(验证它).

在后两个例子中可以毫无困难地指出n个向量的线性无关组. 但对于维数的定义, 还需验证的是, 这些空间没有秩更大的组. 容易想到, 如果运用定理2或它的推

论, 就可以避免去逐个检查各式各样的组.

定义4 设V是域$\mathcal{K}$上的n维向量空间, 任意n个线性无关的向量$\mathbf{e}_1, \cdots, \mathbf{e}_n$构成的组就称为空间$V$的(有限的线性的)**基底**.

有理由认为, 零维空间的基底由向量的空集构成. n维空间的定义已指明了在V中基底的存在性.

下面的定理指出用怎样的方法由给定的基底能实际地建立新的基底.

定理3 设V是$\mathcal{K}$上向量空间, 有基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$. 那么, 有如下论断:

i) 每个向量$\mathbf{v} \in V$都可以唯一地表示成向量$\mathbf{e}_1, \cdots, \mathbf{e}_n$的线性组合形式;

ii) V的任意的线性无关向量组$\mathbf{f}_1, \cdots, \mathbf{f}_s, s \leqslant n$都可以扩充成基底. 特别地, 任意一个向量$\mathbf{v} \neq \mathbf{0}$都可以包含在某一个基底之中.

证明 i) 把给定的向量$\mathbf{v} \in V$和给定的基底放到一起, 按n维空间的定义, 就得到一个线性相关组, 进而得出一个非平凡的关系式

$$\alpha\mathbf{v} + \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0},$$

且系数α应该不等于零. 从而

$$\mathbf{v} = (-\alpha^{-1}\alpha_1)\mathbf{e}_1 + \cdots + (-\alpha^{-1}\alpha_n)\mathbf{e}_n$$

就是基底向量的线性组合.

由两个分解式

$$\beta_1\mathbf{e}_1 + \cdots + \beta_n\mathbf{e}_n = \mathbf{v} = \gamma_1\mathbf{e}_1 + \cdots + \gamma_n\mathbf{e}_n$$

的存在性, 相减后, 我们就能得到

$$(\beta_1 - \gamma_1)\mathbf{e}_1 + \cdots + (\beta_n - \gamma_n)\mathbf{e}_n = \mathbf{0},$$

再由$\mathbf{e}_1, \cdots, \mathbf{e}_n$的线性无关性即可得到所有系数都等于零:

$$\beta_1 - \gamma_1 = \cdots = \beta_n - \gamma_n = 0,$$

也就是$\beta_1 = \gamma_1, \cdots, \beta_n = \gamma_n$. 这刚好建立了分解的唯一性.

ii) 考察向量组

$$\mathbf{f}_1, \cdots, \mathbf{f}_s; \mathbf{e}_1 \cdots \mathbf{e}_n. \tag{1}$$

现在从(1)中选择出一些向量, 使得每个向量都不能由它前面的那些向量线性表示出来. 根据条件, $\mathbf{f}_1, \cdots, \mathbf{f}_s$线性无关, 所以, 其中任意一个都不能由它前面那些向量表示出来. 从而可以假设选择出来的组如下:

$$\mathbf{f}_1, \cdots, \mathbf{f}_s; \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}. \tag{2}$$

任意非平凡的关系式

$$\alpha_1\mathbf{f}_1+\cdots+\alpha_s\mathbf{f}_s+\beta_1\mathbf{e}_{i_1}+\cdots+\beta_t\mathbf{e}_{i_t}=\mathbf{0}$$

若$\beta_k\neq 0$且k是脚码中的最大者, 那么, 在向量组(2)中推知$\mathbf{e}_{i_k}$可由它前面的向量组表示出来, 按其构造方式, 这是不可能的. 另一方面, 由i), V的每个向量均可由基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$线性表出, 从而就能由向量组(1)表出, 进而可由(2)表出. 这样, 线性无关组(2)是极大的, 它就是V的一个基底, 而$\mathbf{e}_{i_1},\cdots,\mathbf{e}_{i_t}$就是找到的补足者. □

在结论ii)的证明中所使用论断, 习惯上, 称之为**施泰尼茨替换原则**. 结论ii)的一个浅易的推论是蕴涵式

$$V_1\subsetneqq V_2\Rightarrow r_1<r_2,$$

其中V_1,V_2是空间V的子空间, 维数分别为r_1和r_2.

说明1　有限维空间V的基底中向量的个数不随基底变化, 有时基底就简单地被认为是空间的子集, 而不在意具体的基底元素(或者基底元素的排列顺序). 应用矩阵形式, 基底元素的标号问题就没有意义了, 这将在下一步弄得更清楚. 指标集的结构整体上由事物的本质确定, 事实上, 指标并不是永远只涉及自然数, 如函数空间$\mathfrak{K}^X$($\mathfrak{K}$是个域, $|X|<\infty$)的Δ函数基底$\{\delta_x|x\in X\}$(见§1之例4)可以很自然地用元素$x\in X$来标记. 而且, 如果X是有限群, 那么, X上取值于$\mathfrak{K}$的所有函数的线性空间$\mathfrak{K}^X$即成为$\mathfrak{K}$上的$|X|$维代数(见§1第2目末尾的定义), 只要令

$$\delta_x*\delta_{x'}=\delta_{xx'},\forall x,x'\in X$$

并将其线性地扩展到所有的函数$f=\sum f(x)\delta_x, g=\sum g(x')\delta_{x'}$上去:

$$f*g=\sum_{x,x'\in X}f(x)g(x')\delta_{xx'}=\sum_{y\in X}\left(\sum_{x\in X}f(x)g(x^{-1}y)\right)\delta_y.$$

该运算被称为**卷积**. 若$X=\{x_1,\cdots,x_n\}$而在V中取用自然数标号的基底$(\Delta_1,\cdots,\Delta_n)$, $\Delta_i=\delta_{x_i}$,那么, 用公式$\Delta_i*\Delta_j=\Delta_k$来确定指标$k$顿时就产生了困难.

3. 坐标. 空间的同构　有了定理3就可以有下面的

定义5　设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是$\mathfrak{K}$上向量空间$V$的一个基底, 由表达式

$$\mathbf{v}=\lambda_1\mathbf{e}_1+\cdots+\lambda_n\mathbf{e}_n,$$

得到的纯量$\lambda_1,\cdots,\lambda_n\in\mathfrak{K}$即称为向量$\mathbf{v}$在给定基底之下的**坐标**.

如果, $\mathbf{x}=\alpha_1\mathbf{e}_1+\cdots+\alpha_n\mathbf{e}_n, \mathbf{y}=\beta_1\mathbf{e}_1+\cdots+\beta_n\mathbf{e}_n$, 那么, $\mathbf{x}+\mathbf{y}=(\alpha_1+\beta_1)\mathbf{e}_1+\cdots+(\alpha_n+\beta_n)\mathbf{e}_n$, 即向量$\mathbf{x},\mathbf{y}$相加时它们的坐标相加. 其次, 同样地, $\lambda\mathbf{x}=\lambda\alpha_1\mathbf{e}_1+\cdots+\lambda\alpha_n\mathbf{e}_n$, 所以, 用$\lambda$乘$\mathbf{x}$时就用同一纯量$\lambda$乘向量$\mathbf{x}$的坐标. 所有坐标均为零的向量恰好是零向量.

如果P_n是以$\mathbb{R}[t]$中次数$\leqslant n-1$的多项式为向量的向量空间, 正如已经指出的, $\mathbf{e}_0 = 1, \mathbf{e}_1 = t, \cdots, \mathbf{e}_{n-1} = t^{n-1}$就是一个基底. 在这个基底之下, 多项式$f(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{n-1} t^{n-1}$的坐标就是它的系数$\alpha_0, \alpha_1, \cdots, \alpha_{n-1}$. 但是, 同样地, $f(t)$可以写成形如

$$f(t) = f(\alpha) + f'(\alpha)(t-\alpha) + \cdots + \frac{f^{(n-1)}(\alpha)}{(n-1)!}(t-\alpha)^{n-1},$$

这就意味着, 在基底$\mathbf{e}'_0 = 1, \mathbf{e}'_1 = t-\alpha, \cdots, \mathbf{e}'_{n-1} = (t-\alpha)^{n-1}$之下, 它有坐标

$$f(\alpha), f'(\alpha), \cdots, \frac{f^{(n-1)}(\alpha)}{(n-1)!}.$$

在$\mathbb{R}^n$中向量$\mathbf{x} = (\alpha_1, \cdots, \alpha_n)$对于基底$\mathbf{e}_1 = (1, 0, \cdots, 0), \mathbf{e}_2 = (0, 1 \cdots, 0), \cdots, \mathbf{e}_n = (0, 0, \cdots, 1)$(从前, 我们记成$E_{(1)}, E_{(2)}, \cdots, E_{(n)}$)的坐标是数$\alpha_1, \alpha_2, \cdots, \alpha_n$(从而就把这个空间称为坐标空间), 但是,$\mathbb{R}^n$有由无穷多个其他的基底构成的集合, 在每个基底之下同一个$\mathbf{x}$的坐标就是一个新的数组.

现在, 看一般状态的这种情形. 设V是域$\mathfrak{K}$上的n维向量空间, $(\mathbf{e}_1, \cdots, \mathbf{e}_n)$, $(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$是它的两个基底. 一个基底的向量用另一个基底的向量表示出来:

$$\begin{aligned} \mathbf{e}'_1 &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + \cdots + a_{n1}\mathbf{e}_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{e}'_n &= a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + \cdots + a_{nn}\mathbf{e}_n. \end{aligned} \tag{3}$$

系数$a_{ij} \in \mathfrak{K}$就决定了一个矩阵

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

称它为由基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$向基底 $(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$的**转换矩阵**. 需要强调的是, 向量$\mathbf{e}'_j$对于基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$的坐标位于矩阵$A$的第$j$列.

设向量$\mathbf{v} \in V$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的坐标是$\lambda_1, \cdots, \lambda_n$,而在一个新基底 $(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$之下的坐标是$\lambda'_1, \cdots, \lambda'_n$, 即

$$\lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n = \mathbf{v} = \lambda'_1\mathbf{e}'_1 + \cdots + \lambda'_n\mathbf{e}'_n.$$

用表达式(3), 以$\mathbf{e}_i$替换$\mathbf{e}'_j$, 得到

$$\begin{aligned} \mathbf{v} &= \lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n \\ &= \lambda'_1(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + \cdots + a_{n1}\mathbf{e}_n) + \cdots + \lambda'_n(a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + \cdots + a_{nn}\mathbf{e}_n), \end{aligned}$$

从而有

$$\begin{aligned}\lambda_1 &= a_{11}\lambda'_1 + a_{12}\lambda'_2 + \cdots + a_{1n}\lambda'_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \lambda_n &= a_{n1}\lambda'_1 + a_{n2}\lambda'_2 + \cdots + a_{nn}\lambda'_n,\end{aligned} \tag{4}$$

或者, 如[BA Ⅰ]第2章所写的那样

$$X = AX' \tag{4'}$$

其中$X = [\lambda_1, \cdots, \lambda_n]$, $X' = [\lambda'_1, \cdots, \lambda'_n]$分别是原有的坐标列和新的坐标列.

公式(4), (4′)借助矩阵A导出的线性变换把向量$\mathbf{v}$的原有的坐标$\lambda_1, \cdots, \lambda_n$用新的坐标$\lambda'_1, \cdots, \lambda'_n$表示出来.

我们一开始就可以用$\mathbf{e}'_1, \cdots, \mathbf{e}'_n$把$\mathbf{e}_1, \cdots, \mathbf{e}_n$表示出来(在$V$中两个基底等价), 从而有公式

$$\lambda'_i = a'_{i1}\lambda_1 + a'_{i2}\lambda_2 + \cdots + a'_{in}\lambda_n, \quad 1 \leqslant i \leqslant n. \tag{5}$$

它们的存在表明, 矩阵A导出的线性变换是可逆的, 即$\det A \neq 0$, 而且(5)变成

$$X' = A^{-1}X, \quad A^{-1} = (a'_{ij}).$$

也就是说, 得到了

定理4　在由矩阵A决定的空间V中从基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$到基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$的转换中, 向量在新的基底之下的坐标可由原有的坐标借助矩阵A^{-1}导出的线性变换表达出来.

要特别注意, 新的基底(上角加了撇"′"的)可按公式(3)由原有基底表示出来, 自然地,按公式(4)可把原来的坐标用新的(上角加了撇"′"的)坐标表示出来(需注意求和顺序), 此时新坐标用原有坐标的表达式则需要复杂的过渡矩阵转换运算. 坐标的运用把向量运算归结为纯量(比方说, 在$\mathbb{R}$上的数t)的作用. 选择合适的坐标系(基底)常常可以从本质上简化计算. 利用基底和坐标系的概念, 我们现在可以从代数学的角度把同一维数的向量空间看成是一样的.

定义6　称域$\mathfrak{K}$上的向量空间V和W是**同构的**, 如果有双射映射$f: V \to W$使得对任意$\alpha, \beta \in \mathfrak{K}, \mathbf{u}, \mathbf{v} \in V$都有

$$f(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v}). \tag{6}$$

换言之, f是V到W的具有补充性质$f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$的一个加法群同构. 同样可以说, f在$\mathfrak{K}$上是线性的, 或者f是$\mathfrak{K}$线性的. 由群同构的定义推知, 映射f^{-1}同样地是W到V的同构. 此外, 同构的合成

$$U \xrightarrow{g} V \xrightarrow{f} W$$

$f \circ g : U \to W$还是同构, 可以直接看出, 维数是同构不变的: 如果$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是$V$的一个基底, 那么, $(f(\mathbf{e}_1), \cdots, f(\mathbf{e}_n))$就是$W$的一个基底. 反之亦然.

下面的定理表明, 没有其他的同构不变量.

定理5 $\mathcal{K}$上所有的同一维数n的向量空间都同构. 更精确地, 它们都同构于坐标空间$\mathcal{K}^n$.

证明 设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是$n$维空间的任意一个基底. 任意向量 $\mathbf{x} = \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n$的坐标$\alpha_1, \cdots, \alpha_n$是唯一确定的, 所以,$V$中向量和$\mathcal{K}^n$中的向量之间的对应

$$f : \mathbf{x} \mapsto (\alpha_1, \cdots, \alpha_n)$$

是个双射. 如果$\mathbf{y} = \beta_1\mathbf{e}_1 + \cdots + \beta_n\mathbf{e}_n$, 那么

$$\alpha\mathbf{x} + \beta\mathbf{y} = (\alpha\alpha_1 + \beta\beta_1)\mathbf{e}_1 + \cdots + (\alpha\alpha_n + \beta\beta_n)\mathbf{e}_n.$$

从而有

$$\begin{aligned} f(\alpha\mathbf{x} + \beta\mathbf{y}) &= (\alpha\alpha_1 + \beta\beta_1, \cdots, \alpha\alpha_n + \beta\beta_n) \\ &= \alpha(\alpha_1, \cdots, \alpha_n) + \beta(\beta_1, \cdots, \beta_n) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}), \end{aligned}$$

这就是同构性质的表达式. □

定理的存在性表明, 只要先选择好V的一个基底就可以过渡到$\mathcal{K}^n$. 但是, 仅仅限于在$\mathcal{K}^n$中研究线性问题至少是个不方便的限制, 因为真正的目标是得到与基底选择无关的结论. 此外, 转化到$\mathcal{K}^n$上去将损失掉很多空间的直观特性, 如一般3维空间, 多项式空间, 等等.

注意 在两个向量空间V, W中, 如果只有一个唯一确定的同构, 那么, 只有两种情形: 1)V=W=$\{\mathbf{0}\}$; 2)dim V=1=dim W, 而且$\mathcal{K}$是个二元域(试证明之).

有时, 两个向量空间确定的某个同构与人的意愿, 比如说在V中与在W中的基底选择没有关系, 这样的同构将称为**标准的**或者**自然的**, 用以区别于所有其余的, 即"偶然的".下一节我们将会遇到自然同构的有特性的例子.

4. 子空间的交集与和 我们把众所周知的集合理论中的交和并运算应用到子空间上. 两个子空间$U_1, U_2 \subset V$的交集$U_1 \cap U_2$, 显然地, 是个子空间, 同样可以照搬到任意子空间族$\{U_i | i \in I\}$的交集$U = \bigcap_{i \in I} U_i$上($U$可能是零空间). 实际上, 零向量属于所有的$U_i$, 从而属于$U$, 因而$U$非空. 其次, 如果$\mathbf{x}$, $\mathbf{y} \in U$, 那么, 它们的任意的线性组合$\alpha\mathbf{x} + \beta\mathbf{y}$必属于所有的$U_i$, 从而有$\alpha\mathbf{x} + \beta\mathbf{y} \in U$.

注意, 两个子空间的并集$U_1 \cup U_2$未必是子空间. 例如, 若$\mathbf{e}_1, \mathbf{e}_2$是$V$中任意两个线性无关的向量且$U_1 = \langle \mathbf{e}_1 \rangle, U_2 = \langle \mathbf{e}_2 \rangle$,那么, $U_1 \cup U_2$不包含$\mathbf{e}_1 + \mathbf{e}_2$.

容易看出, V的包含U_1和U_2的最小的子空间显然是

$$U = \{\mathbf{u}_1 + \mathbf{u}_2 | \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2\}.$$

这个子空间称之为U_1, U_2的和, 且用U_1+U_2代表, 很清楚, $U_1+U_2 = U_2+U_1$, 还有, $U_1+U_2 = U_2$则必要而只要$U_1 \subset U_2$. 类似地可定义任意有限个向量子空间$U_1, \cdots, U_m$的和, 意即, $U_1 + \cdots + U_m$就是包含所有取自于$U_i, 1 \leqslant i \leqslant m$的向量的各种可能的线性组合的最小子空间. 此时,任何一种排列括号的方式都一样, 因为$U_i + (U_j + U_k) = (U_i + U_j) + U_k$.

如果A, B是3维物理空间的两个图形,其交$A \cap B$可能非空, 而它们的体积是vol(A), vol(B), 那么, 关系式

$$\operatorname{vol}(A \cup B) = \operatorname{vol}(A) + \operatorname{vol}(B) - \operatorname{vol}(A \cap B)$$

成立.

与其类似的空间的情形可表达成

定理6 设U和W是向量空间V的有限维子空间. 那么[1],

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W). \tag{7}$$

证明 我们设

$$\dim U = k, \dim W = l, \dim(U \cap W) = m.$$

因为$U \cap W \subset U, W$, 故$m \leqslant k$, $m \leqslant l$. 在$U \cap W$中选择那样的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m)$, 根据定理3它一方面可以扩充成$U$的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m; \mathbf{a}_1, \cdots, \mathbf{a}_{k-m})$, 而另一方面它又可以扩充成$W$的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m; \mathbf{b}_1, \cdots, \mathbf{b}_{l-m})$. 和$U + W$的每一个向量必形如$\mathbf{u}+\mathbf{w}$, 其中$\mathbf{u} \in U$, $\mathbf{w} \in W$, 而这意味着

$$U + W = \langle \mathbf{e}_1, \cdots, \mathbf{e}_m; \mathbf{a}_1, \cdots, \mathbf{a}_{k-m}; \mathbf{b}_1, \cdots, \mathbf{b}_{l-m} \rangle.$$

如果, 我们能够证明向量组

$$\mathbf{e}_1, \cdots, \mathbf{e}_m; \mathbf{a}_1, \cdots, \mathbf{a}_{k-m}; \mathbf{b}_1, \cdots, \mathbf{b}_{l-m}$$

是线性无关的, 那么就有了与(7)一致的关系式

$$\dim(U + W) = m + (k - m) + (l - m) = k + l - m,$$

于是证明就完成了. 假设不然, 会

$$\sum_{s=1}^{m} \gamma_s \mathbf{e}_s + \sum_{i=1}^{k-m} \alpha_i \mathbf{a}_i + \sum_{j=1}^{l-m} \beta_j \mathbf{b}_j = \mathbf{0} \tag{*}$$

1) 公式(7)与格拉斯曼(G.H.Grassmann, 1809—1877)的名字相联.

是个非平凡的关系式. 于是, 我们有

$$\sum_{s=1}^{m}\gamma_s\mathbf{e}_s+\sum_{i=1}^{k-m}\alpha_i\mathbf{a}_i=-\sum_{j=1}^{l-m}\beta_j\mathbf{b}_j,$$

其中, 等式的左侧是U中的元素, 而右侧是W中的元素. 这表明, 它是$U\cap W$的元素, 我们可以记之为$-\sum\limits_{j=1}^{l-m}\beta_j\mathbf{b}_j=\sum\limits_{s=1}^{m}\delta_s\mathbf{e}_s$, 或者

$$\sum_{s=1}^{m}\delta_s\mathbf{e}_s+\sum_{j=1}^{l-m}\beta_j\mathbf{b}_j=\mathbf{0}.$$

然而, 子空间W的基底$\{\mathbf{e}_1,\cdots,\mathbf{e}_m;\mathbf{b}_1,\cdots,\mathbf{b}_{l-m}\}$的线性关系式应该是平凡的, 特别地, 有$\beta_1=\cdots=\beta_{l-m}=0$. 进而, 现在已经变成子空间$U$的基底$\{\mathbf{e}_1,\cdots,\mathbf{e}_m;\mathbf{a}_1,\cdots,\mathbf{a}_{k-m}\}$的关系式$(*)$也应该是平凡的, $\gamma_1=\cdots=\gamma_m=\alpha_1=\cdots=\alpha_{k-m}=0$. 我们得到了所需要的矛盾. □

因为, 和$U+W$的维数不能超过规模大的空间V的维数, 所以, 在定理6的基础上, 经常可以得到关于子空间的交的非平凡性的结论. 例如, 3维空间的两个平面或者5维空间中的两个3维子空间都必然包含公共直线, 因为这两种情形都有

$$\dim U+\dim W>\dim V.$$

关于可利用的术语, 我们有如下的

说明2 在n维向量空间V中存在所有各个小维数的子空间, 容易相信, 可以在V中嵌入子空间链

$$0\subset V_1\subset V_2\subset\cdots\subset V_{n-1}\subset V_n=V=\langle\mathbf{e}_1,\cdots,\mathbf{e}_n\rangle,$$

其中$V_i=\langle\mathbf{e}_1,\cdots,\mathbf{e}_i\rangle$. 1 维的向量空间被称为**直线**, 2 维向量空间是**平面**, 当$k\geqslant 3$ 时, 就称是k **维平面**. 设U 是向量空间V 的子空间. 差

$$\operatorname{codim}U=\dim V-\dim U$$

被称为子空间U的**余维数**. 余维数为1的子空间**被称为超平面**. 超平面的概念是相对的, 直线是2维向量空间W的超平面, 而W却可能看成是更大维数的向量空间V的一个平面.

5. 直和 在非零的线性子空间的和

$$U=U_1+U_2+\cdots+U_m \tag{8}$$

中, 任意向量$\mathbf{u}\in U$都可以写成形如

$$\mathbf{u}=\mathbf{u}_1+\mathbf{u}_2+\cdots+\mathbf{u}_m,\quad \mathbf{u}_i\in U_i \tag{9}$$

一般地说, 表示方法不是唯一的.

定义7 如果每个向量$\mathbf{u} \in U$都能唯一地表成(9)的形式, 那么, 就称(8)是**直和**而且用

$$U = U_1 \oplus \cdots \oplus U_m$$

表示.

和(8)是直和, 只要对零向量的情形(9)有唯一的表示. 也就是说,

$$\mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_m \Rightarrow \mathbf{u}_1 = \mathbf{0}, \mathbf{u}_2 = \mathbf{0}, \cdots, \mathbf{u}_m = \mathbf{0}.$$

事实上, 如果这个较弱的条件成立, 那么, 由展开式

$$\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_m = \mathbf{u} = \mathbf{u}_1' + \mathbf{u}_2' + \cdots + \mathbf{u}_m'$$

就可以得到$\mathbf{0} = (\mathbf{u}_1 - \mathbf{u}_1') + (\mathbf{u}_2 - \mathbf{u}_2') + \cdots + (\mathbf{u}_m - \mathbf{u}_m')$, 其中$\mathbf{u}_i - \mathbf{u}_i' \in U_i$. 按假定, $\mathbf{u}_i - \mathbf{u}_i' = \mathbf{0}$, $1 \leqslant i \leqslant m$ 或者$\mathbf{u}_1 = \mathbf{u}_1', \mathbf{u}_2 = \mathbf{u}_2', \cdots, \mathbf{u}_m = \mathbf{u}_m'$, 即, 直和的展开性质成立.

约定一表达方法

$$U_1 + \cdots + \hat{U}_i + \cdots + U_m = U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m.$$

定理7 和$U = U_1 + U_2 + \cdots + U_m$是直和的充要条件是

$$U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m) = \mathbf{0} \tag{10}$$

对每个$i = 1, 2, \cdots, m$都成立.

证明 设这个和是个直和, 看任意向量$\mathbf{x} \in U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m)$, 其中指标$i$是固定的. 这样, $\mathbf{x} = \mathbf{u}_1 + \cdots + \hat{\mathbf{u}}_i + \cdots + \mathbf{u}_m$, 于是对零向量我们可得两个等式

$$\begin{aligned}\mathbf{0} + \cdots + \mathbf{0} + \mathbf{0} + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}& \\ &= \mathbf{u}_1 + \cdots + \mathbf{u}_{i-1} + (-\mathbf{x}) + \mathbf{u}_{i+1} + \cdots + \mathbf{u}_m.\end{aligned}$$

因为这是个直和, 所以这两个分解应该重合. 特别地, $-\mathbf{x} = \mathbf{0}$, 从而(10)式成立.

反过来, 设(10)为真, 去证明零向量分解的唯一性(正如我们已经知道的, 这就足以使和为直和). 事实上, 可从任意一个分解式

$$\mathbf{0} = \mathbf{a}_1 + \cdots + \mathbf{a}_i + \cdots + \mathbf{a}_m$$

出发, 于是对任意$i = 1, 2, \cdots, m$有

$$-\mathbf{a}_i = \mathbf{a}_1 + \cdots + \mathbf{a}_{i-1} + \mathbf{a}_{i+1} + \cdots + \mathbf{a}_m \in U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m) = \mathbf{0}.$$

从而, $\mathbf{a}_i = \mathbf{0}$. □

在m=2的情形, 定理7变成特别简洁的公式: 和$U=U_1+U_2$是直和$\Leftrightarrow U_1\cap U_2=0$. 特别地, 援引关系式(7), 就得到$\dim U=\dim U_1+\dim U_2$. 这个性质的一般性可表述成

定理8　和$U=U_1+U_2+\cdots+U_m$是直和, 当且仅当,

$$\dim U=\sum_{i=1}^{m}\dim U_i. \tag{11}$$

证明　可用对m的归纳法实现之. 当m=2时, 命题的真确性已经在前面指明了. 而对于任意m的情形可采用定理6和定理7. 就是说. 如果和是直和, 那么$U_1+\cdots+\hat{U}_i+\cdots+U_m$也是直和, 于是, 有

$$\begin{aligned}\dim U=&\dim U_i+\dim(U_1+\cdots+\hat{U}_i+\cdots+U_m)-\\&\dim U_i\cap(U_1+\cdots+\hat{U}_i+\cdots+U_m)\\=&\dim U_i+(\dim U_1+\cdots+\dim\hat{U}_i+\cdots+\dim U_m)-0=\sum_{i=1}^{m}\dim U_i.\end{aligned}$$

反过来, 如果公式(11)成立, 那么所有子空间U_i的基底合并起来就是U的一个基底, 进而可知和是直和. □

略加变化, 可得

定理9　对n维空间V的任意一个m维子空间U, 都有一个$n-m$维的子空间W使得$V=U\oplus W$(称U和W是**互补**的子空间).

证明　设$(\mathbf{a}_1,\cdots,\mathbf{a}_m)$是$U$的任意一个基底, 将其扩展成$V$的一个基底$(\mathbf{a}_1,\cdots,\mathbf{a}_m;\mathbf{b}_1,\cdots,\mathbf{b}_{n-m})$(用定理3). 令$W=\langle\mathbf{b}_1,\cdots,\mathbf{b}_{n-m}\rangle$, 立刻就得结论. □

到现在为止, 我们所关注的直和都是在一个固定的向量空间V中, 这种直和通常称为**内直和**. 但是, 有时有必要考察同一个域$\mathcal{K}$上两个向量空间的**外直和**, 而这两个空间并不预先作为某个空间的子空间. 这时, $U\oplus W$应该被理解为所有有序对$(\mathbf{u},\mathbf{w})$, $\mathbf{u}\in U,\mathbf{w}\in W$的总体$V=U\times W$. V中加法和纯量乘运算按公式

$$\alpha(\mathbf{u},\mathbf{w})+\beta(\mathbf{u}',\mathbf{w}')=(\alpha\mathbf{u}+\beta\mathbf{u}',\alpha\mathbf{w}+\beta\mathbf{w}').$$

实现. 这类似于用两个坐标轴构作一个平面.

所有的向量$(\mathbf{u},\ \mathbf{0})$构成$V$的一个子空间$\widetilde{U}$, 它同构于$U$, 而向量$(\mathbf{0},\ \mathbf{w})$构成同构于$W$的子空间$\widetilde{W}$. 这里的同构$(\mathbf{u},\mathbf{0})\mapsto\mathbf{u},(\mathbf{0},\mathbf{w})\mapsto\mathbf{w}$是显然的; 同时, 可以记成:

$$\underbrace{U\oplus W}_{\text{外直和}}=V=\underbrace{\widetilde{U}\oplus\widetilde{W}}_{\text{内直和}},$$

因为我们已经把$\widetilde{U}\oplus\widetilde{W}$看作向量空间$V$的直和. 进一步要讨论的大多是关于内直和的, 因此, 所有的细节都可以省略掉.

6. 商空间　一般来说, 对一个给定的子空间$L \subset V$会有很多的补子空间$W \subset V$使得$L \oplus W = V$, 而且, 所有这些补子空间都同构于同一个向量子空间, 它由V和L用绝对不变量形成, 而不依赖任何个人意愿.

把V和L看成阿贝尔加法群, 集合

$$\mathbf{x} + L = \{\mathbf{x} + \mathbf{y} | \mathbf{y} \in L\}$$

说成是V对于L的一个**陪集**, 向量$\mathbf{x}$是这个陪集的一个代表元. 如果$\mathbf{0} \neq \mathbf{z} \in (\mathbf{x} + L) \cap (\mathbf{x}' + L)$. 那么, $\mathbf{x}+\mathbf{y} = \mathbf{x}'+\mathbf{y}' = \mathbf{z}$且$\mathbf{x}+L = \mathbf{x}'+L = \mathbf{z}+L$. 所以, 两个陪集或者不相交或者是重合的. 对固定的L, 设$\mathbf{x} := \mathbf{x} + L$, 每个向量$\mathbf{v} \in V$必落入这样的一个陪集, 而且, 若$\overline{V} = V/L$是所有$V$对于$L$的陪集的集合. 那么, 在$\overline{V}$上按照$\overline{\mathbf{x}} + \overline{\mathbf{x}'} = \overline{\mathbf{x} + \mathbf{x}'}$规则就建立了$\overline{V}$上的阿贝尔群结构. 交换性和结合性可直接验证. 显然, $\bar{\mathbf{0}} = L$是这个阿贝尔群的零元: $\overline{\mathbf{x}} + \bar{\mathbf{0}} = \overline{\mathbf{x} + \mathbf{0}} = \overline{\mathbf{x}}$. 其次, $-\overline{\mathbf{x}} = \overline{-\mathbf{x}}$.

令$\lambda\overline{\mathbf{x}} = \overline{\lambda\mathbf{x}}$, 也就是, $\lambda(\mathbf{x} + L) = \lambda\mathbf{x} + L, \quad \forall\lambda \in \mathfrak{K}$, 我们容易相信, §1中的所有公理$\text{B}\Pi_1$—$\text{B}\Pi_8$都得到满足. 例如

$$1 \cdot (\mathbf{x} + L) = 1 \cdot \mathbf{x} + L = \mathbf{x} + L,$$
$$\alpha(\beta(\mathbf{x} + L)) = \alpha(\beta\mathbf{x} + L) = \alpha\beta\mathbf{x} + L = (\alpha\beta)(\mathbf{x} + L).$$

这样一来, $\overline{V} = V/L$以自然的方式获得了向量空间的结构, 称它是空间V对子空间L的**商空间**(或者说是模L的商空间). 在陪集中我们考察由同余式

$$\mathbf{x} \equiv \mathbf{x}' (\text{mod} L) \Leftrightarrow \mathbf{x} - \mathbf{x}' \in L,$$

确定的等价类, 这些都是套话.

例6　设$V = \mathbb{R}^2$是坐标平面, 而L是x轴. 经过O点引出的任意一条非水平线都可充当L的余子空间M(图2).

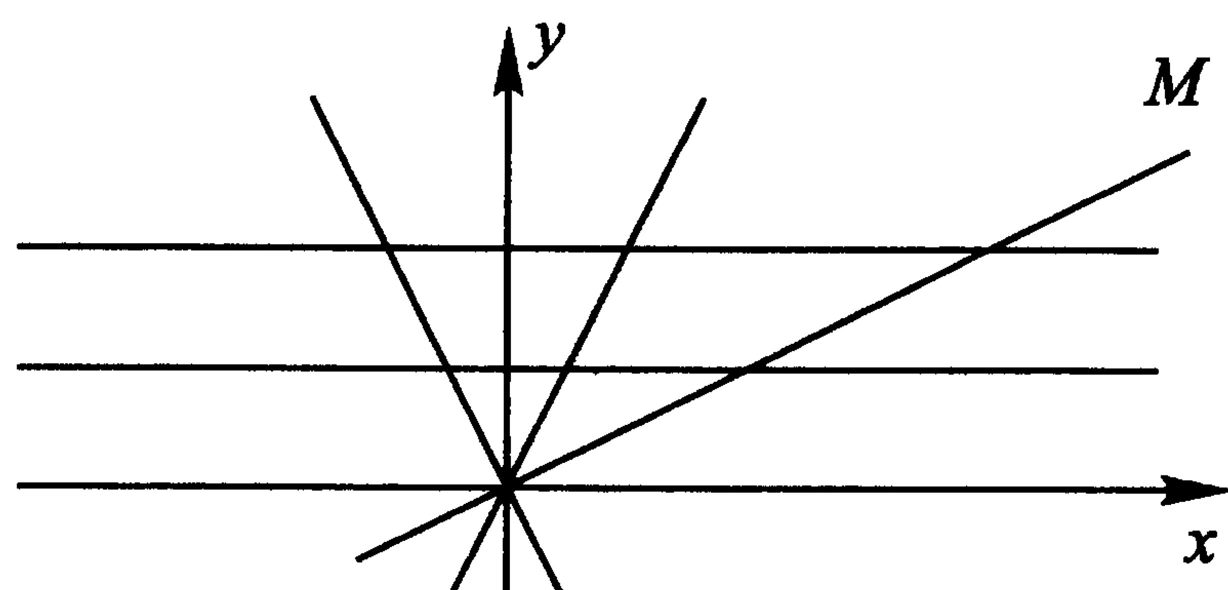


图2

余集M与每个平行于x轴的直线只相交于一点, 因此, M可随所有这些直线的集合变化, 这个集合恰好就是V/L.

定理10　设$V = L \oplus M$是个直和, $L, M \subset V$. 那么, 映射$f : \mathbf{u} \mapsto \mathbf{u} + L \ (\mathbf{u} \in M)$是$M$与$V/L$之间的同构.

证明 事实上, f是个线性映射, 因为

$$f(\alpha\mathbf{u}+\beta\mathbf{v})=\alpha\mathbf{u}+\beta\mathbf{v}+L=\alpha(\mathbf{u}+L)+\beta(\mathbf{v}+L)=\alpha f(\mathbf{u})+\beta f(\mathbf{v}).$$

设$\mathbf{v}+L$是V/L的一个任意元素, 按条件, $\mathbf{v}=\mathbf{x}+\mathbf{y},\mathbf{x}\in L,\mathbf{y}\in M$, 所以,

$$\mathbf{v}+L=\mathbf{x}+\mathbf{y}+L=(\mathbf{x}+L)+(\mathbf{y}+L)=L+(\mathbf{y}+L)=\mathbf{y}+L=f(\mathbf{y}).$$

这表明f是个满射. 其次, 如果$\mathbf{u}\in\mathrm{Ker}f$, 那么$\mathbf{u}+L=L$, 因此$\mathbf{u}\in L$. 但$\mathbf{u}\in M$, 而$L\cap M=\mathbf{0}$,所以, $\mathbf{u}=\mathbf{0}$, 故$\mathrm{Ker}\, f=\mathbf{0}$. 可见, f是个双射. □

推论 设L是V的任意一个子空间. 那么,

$$\dim V/L=\dim V-\dim L.$$

换言之, $\dim V/L=\mathrm{codim}_V L$.

证明 按照定理9可以找到一个子空间$W\subset V$使得$V=L\oplus M$. 从而$\dim M=\dim V-\dim L$. 再按照已经证明了的定理10, 这个子空间M必然同构于商空间V/L. □

习 题

1. 在q元域$\mathbb{F}_q$上的n维向量空间V中, 对于$1\leqslant k\leqslant n$, 有多少个k维的子空间?

2. 试指出由下列各种n阶实方阵构成的空间的维数:

(1) 对称矩阵; (2)斜对称矩阵; (3)迹为零的矩阵.

3. 由所有的一个变元t的次数$\leqslant n$的满足条件$f(1)=0$的多项式组成的空间的维数是多少? 找出这个空间的一个基底来.

4. 和$U=U_1+U_2+\cdots+U_m$是直和的充要条件是

$$(U_1+\cdots+U_{i-1})\cap U_i=0,\quad 1<i\leqslant m.$$

5. 找出空间P_n 的基底$(1,t,\cdots,t^{n-1})$ 向同一空间的基底$(1,(t-\alpha),\cdots,(t-\alpha)^{n-1})$ 的转换矩阵.

6. 设θ是$\mathbb{Q}$上不可约多项式$f\in\mathbb{Q}\,[t]$的一个复根. 求出空间$\mathbb{Q}\,[\theta]=\langle 1,\theta,\cdots,\theta^k,\cdots\rangle_{\mathbb{Q}}$在$\mathbb{Q}$上的维数.

7. 证明, 直和不满足消去律, 也就是说, 由等式$U\oplus W_1=U\oplus W_2$, 一般说来, 不能得到$W_1=W_2$, 尽管等式有一相同的被加项.

8. 查明下列商空间$\mathfrak{R}[t]/L$是否是有限维的:

(1) L是所有次数$\leqslant n-1$的t多项式P_n构成的子空间;

(2) L是所有可以被t^n整除的关于t的多项式构成的子空间;

(3) L是所有的关于t^2的多项式构成的子空间.

9. 证明下面与格拉斯曼公式类似的等式

$$\mathrm{codim}(U+W)+\mathrm{codim}(U\cap W)=\mathrm{codim}\ U+\mathrm{codim}\ W$$

(U与W是向量空间V的有限余维子空间, V不一定是有限维的).

10. 遵循§1的例8中的术语, 我们挑选出显而易见的半幻矩阵

$$0,\quad E=\begin{pmatrix}1&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix},\quad D=\begin{pmatrix}&&&1\\&&1&\\&\cdot^{\cdot^{\cdot}}&&\\1&&&\end{pmatrix}, S=\begin{pmatrix}1&1&\cdots&1\\1&1&\cdots&1\\\cdots&\cdots&\cdots&\cdots\\1&1&\cdots&1\end{pmatrix}.$$

问题在于: 有怎样的维数$\dim \mathrm{SMag}_n(\mathbb{Q})$和$\dim \mathrm{Mag}_n(\mathbb{Q})$?显然, $\mathrm{SMag}_2(\mathbb{Q})=\langle E,D\rangle_{\mathbb{Q}}$. 此时, $S=E+D$是唯一的(精确到魔幻矩阵有理因子). 当n=3时, 可指出一个略明显的魔幻矩阵

$$A=\begin{pmatrix}1&2&0\\0&1&2\\2&0&1\end{pmatrix}.$$

计算出n=3和n=4时上面提到的维数.

11. 证明直和分解式

$$\mathrm{SMag}_n(\mathbb{Q})=\mathrm{Mag}_n(\mathbb{Q})\oplus\mathbb{Q}E\oplus\mathbb{Q}D.$$

12. 设$V_1,\cdots,V_k$是n维向量空间V的子空间. 证明, 如果$\dim V_1+\cdots+\dim V_k>n(k-1)$, 那么, $\bigcap\limits_{i=1}^{k}U_i\neq\mathbf{0}$(由公式(7)导出的论断的直接推广).

§3 对偶空间

1. 线性函数 域$\mathfrak{K}$上任意一个有限维的向量空间V都可以与另外一个和V有特殊对偶关系的向量空间相对应. 为此目的, 我们引进

定义1 映射$f:V\to\mathfrak{K}$, 有性质

$$f(\alpha\mathbf{x}+\beta\mathbf{y})=\alpha f(\mathbf{x})+\beta f(\mathbf{y}),\quad \forall\alpha,\beta\in\mathfrak{K}\quad \mathbf{x},\mathbf{y}\in V,$$

即称f为是V上的**线性函数**(线性形式, 线性型或线性泛函; 后者多应用于无穷维空间理论).

在V中选择任意一个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$, 把线性函数$f$作用到向量$\mathbf{x}=\lambda_1\mathbf{e}_1+\cdots+\lambda_n\mathbf{e}_n$上并记成

$$f(\mathbf{x})=\lambda_1\beta_1+\cdots+\lambda_n\beta_n \tag{1}$$

形式, 其中$\beta_i=f(\mathbf{e}_i)$是纯量, 只与基底的选择有关. 反过来, 可以直接看出, 对于一个给定的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$而言, 任意纯量$\beta_i\in\mathfrak{K}, i=1,2,\cdots,n$, 有而且只有一个线性函数符合要求.

特别需要注意, 与线性函数的定义等价的关系式

$$f(\mathbf{x}+\mathbf{y})=f(\mathbf{x})+f(\mathbf{y}),\quad f(\lambda\mathbf{x})=\lambda f(\mathbf{x})$$

并没有涉及任何有关基底选择的事情. 也就是说, 线性函数的定义是不变的(与基底的选择无关). 把线性函数值表达成(1), 我们应当知道一个规则, 即在一个基底向另一个基底转换时, 系数$\beta_i = f(\mathbf{e}_i)$如何变化. 设

$$\langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle = V = \langle \mathbf{e}'_1, \cdots, \mathbf{e}'_n \rangle,$$
$$\mathbf{e}'_j = a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \cdots + a_{nj}\mathbf{e}_n \quad j = 1, 2, \cdots, n,$$

是由原来的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$向新基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$转换时的公式. 现在如果

$$\lambda_1\beta_1 + \cdots + \lambda_n\beta_n = f(\mathbf{v}) = \lambda'_1\beta'_1 + \cdots + \lambda'_n\beta'_n,$$

其中$\lambda_1, \cdots, \lambda_n$和$\lambda'_1, \cdots, \lambda'_n$是向量$\mathbf{v} \in V$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$和基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$之下分别对应的坐标, 那么, 容易看出,

$$\begin{aligned} \beta'_j &= f(\mathbf{e}'_j) = f(a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \cdots + a_{nj}\mathbf{e}_n) \\ &= a_{1j}f(\mathbf{e}_1) + a_{2j}f(\mathbf{e}_2) + \cdots + a_{nj}f(\mathbf{e}_n) \\ &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n. \end{aligned} \tag{2}$$

可见, 在基底转换时, 基底向量的系数和线性形式的系数按照相同的公式变化, 也就是一致的, 或者还可以说是同步的.

2. 对偶空间与对偶基底 有了V上的线性函数f, g就可以研究用任意$\alpha, \beta \in \mathfrak{K}$构作出来的它们的线性组合$\alpha f + \beta g$, 令

$$(\alpha f + \beta g)(x) \xlongequal{\text{def}} \alpha f(x) + \beta g(x).$$

可直接验证, $\alpha f + \beta g$也是个线性函数. 从而, 下面的定义有意义.

定义2 对于刚引入的加法和纯量乘运算, 所有的线性函数作成一个**对偶**于(或**共轭**于)V的向量空间$V^* = \mathcal{L}(V, \mathfrak{K})$.

说明 在同时要处理向量空间V和V^*的时候, 称V^*的元素为**共变向量**(或**余向量**), 而称V的元素为**反变向量**, 在我们将于第6章中加以关注的一般的张量理论中, 余向量与(1, 0)型张量(1阶的**共变张量**)有关, 而向量与(0, 1)型张量(1阶**反变张量**)有关. 换用新的术语来说. 相应于基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ 和基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$, 且

$$\mathbf{e}'_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i, \quad j = 1, \cdots, n,$$

纯量$\beta_1, \cdots, \beta_n$上角带小撇与不带撇的组(它们之间用公式(2)联系起来)的对应称为 $(1, 0)$型张量. 对于用相应的基底下的组$(\lambda_1, \cdots, \lambda_n)$, $(\lambda'_1, \cdots, \lambda'_n)$决定(0, 1)型张量, 同样, 可用§2的公式(4)表示出来. 前缀"共"和"反"在数学中会经常遇到(最一般的理解是共变函子和反变函子), 它们的意义总是大致相同且在某种程度上用已研究过的最简单的例子来说明. 通用的张量记将在晚些时候引入.

我们感到, 在空间V的给定基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下, 在线性函数和$n$元纯量组之间有相互单值的对应$\Phi: f\mapsto(\beta_1,\cdots,\beta_n)$. 我们可以把这个纯量组与坐标向量空间$\mathfrak{K}^n$的向量等同起来, 而且可以注意到, 如果$f\mapsto(\beta_1,\cdots,\beta_n)$, $g\mapsto(\gamma_1,\cdots,\gamma_n)$, 那么

$$f+g\mapsto(\beta_1+\gamma_1,\cdots,\beta_n+\gamma_n),\quad \lambda f\mapsto(\lambda\beta_1,\cdots,\lambda\beta_n).$$

这样一来, Φ是向量空间V^*和$\mathfrak{K}^n$之间的同构, 特别地, $\dim V^*=\dim\mathfrak{K}^n=n$.

取$\beta_j=0$, 当$j\neq i$时; $\beta_i=1$, 且设

$$e^i(\mathbf{e}_j)=\delta_{ij},\quad j=1,\cdots n.$$

我们就定义了线性函数$e^i\in V^*$:

$$e^i\left(\sum\lambda_j\mathbf{e}_j\right)=\sum\lambda_j e^i(\mathbf{e}_j)=\sum\lambda_j\beta_j=\lambda_i.$$

函数$e^1,\cdots,e^n$ 显然是线性无关的, 因为它们在$\mathfrak{K}^n$ 中对应线性无关的行向量 $(0,\cdots,1,\cdots,0)$.

这就证明了

定理1 设V是域$\mathfrak{K}$上的n维向量空间. 那么, 对偶空间V^*同样也是n维的. 如果$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是$V$的一个基底, 而线性函数$e^1,\cdots,e^n$使得

$$e^i(\mathbf{e}_j)=\delta_{ij}=\begin{cases}1, i=j,\\0, i\neq j.\end{cases}$$

那么, $(e^1,\cdots,e^n)$就是V^*的一个基底.

定义3 按照定理1的叙述给出的V^*的基底$(e^1,\cdots,e^n)$称为是相对于V的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$的**对偶基底**(或相互的对偶基底).

对偶于V的空间V^*及对偶基底$(e^1,\cdots,e^n)$, $(\mathbf{e}_1,\cdots,\mathbf{e}_n)$的称谓本身就在$V$和$V^*$之间建立了"双边对称"的关系, 它的性质将随着新的概念的引入而逐步显露出来. 我们暂时约定, 用记法$(f,\mathbf{x})$代替$f(\mathbf{x})$, 暗示这是一个纯量积, 但来自两个不同的空间, 这本身就决定了一个对于每个变量来说都是线性的映射$V^*\times V\to\mathfrak{K}$:

$$(\alpha f+\beta g,\mathbf{x})=\alpha(f,\mathbf{x})+\beta(g,\mathbf{x}),\quad (f,\alpha\mathbf{x}+\beta\mathbf{y})=\alpha(f,\mathbf{x})+\beta(f,\mathbf{y})\tag{3}$$

具有这种性质的映射$V\times W\to\mathfrak{K}$通常称为是**双线性**的, 同时称V和W之间是配对的. 我们讨论过的V^*和V之间的这种配对说成是**规范**的.

利用对偶基并用它们把元素表示出来:

$$\mathbf{x}=\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\cdots+\alpha_n\mathbf{e}_n,\quad f=\beta_1e^1+\beta_2e^2+\cdots+\beta_ne^n,$$

容易算出值

$$f(\mathbf{x})=(f,\mathbf{x})=\alpha_1\beta_1+\alpha_2\beta_2+\cdots+\alpha_n\beta_n.$$

另一方面，可以计算出$\mathbf{x}$对于基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$的坐标$\alpha_k$和余向量(线性函数)$f$对于基底$(e^1,\cdots,e^n)$的坐标公式

$$\alpha_k=(e^k,\mathbf{x}),\quad \beta_k=(f,\mathbf{e}_k) \tag{4}$$

事实上，

$$(e^k,\mathbf{x})=(e^k,\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\cdots+\alpha_n\mathbf{e}_n)=\sum_i\alpha_i(e^k,\mathbf{e}_i)=\alpha_k,$$

$$(f,\mathbf{e}_k)=\left(\sum_i\beta_i e^i,\mathbf{e}_k\right)=\sum_i\beta_i(e^i,\mathbf{e}_k)=\beta_k.$$

例 设$V=P_n=\langle 1,t,\cdots,t^{n-1}\rangle$是次数$\leqslant n<1$的$n$维实多项式空间. 映射$f_\lambda:\varphi\mapsto\varphi(\lambda)$. 在每个点$\lambda\in\mathbb{R}$, 多项式$\varphi(t)=\varphi_0+\varphi_1t+\cdots+\varphi_{n-1}t^{n-1}$都对应它在该点的值, f_λ显然是线性的. 变动λ, 就可以得到对偶空间V^*的基底, 可以方便地引入函数$f:V\to\mathbb{R},\quad f(\varphi)=\mu\varphi^{(k)}(\lambda)$, 其中$\varphi^{(k)}$是多项式$\varphi$的$k$阶导数, 而$\mu$, λ是固定数. 因为

$$f(\alpha\varphi+\beta\psi)=\mu(\alpha\varphi+\beta\psi)^{(k)}(\lambda)=\mu(\alpha\varphi^{(k)}(\lambda)+\beta\psi^{(k)}(\lambda))=\alpha f(\varphi)+\beta f(\psi).$$

所以, $f\in V^*$. 特别地, 线性函数

$$e^k:\varphi\mapsto\frac{\varphi^{(k)}(0)}{k!}=\varphi_k,\quad k=0,1,\cdots,n-1.$$

构成V^*的一个基底, 它对偶于$1,t,\cdots,t^{n-1}$.而对偶于基底$1,t-\lambda,\cdots,(t-\lambda)^{n-1}$可选择函数$\varphi\mapsto\varphi^{(k)}(\lambda)/k!(k=0,1,\cdots,n-1)$组成基底. 这个关系让我们想起函数的泰勒分解系数.

3. 自反性 直接把定理1和§2的定理5联系到一起即导出结论: 至少在$\dim V<\infty$的情形有同构$V^*\cong V$. 同理, 应有V^*与V^{**}=$(V^*)^*$之间的同构. 按定义V^{**}是V^*的对偶空间, 也就是V^*的线性函数空间. 乍一看, 似乎难以用合理的方式用原有空间V的术语解释V^{**}的元素. V^{**}有与V的自然对应, 如下所示

定理2 *存在自然同构$\varepsilon:V\to V^{**}$, 它由公式*

$$\varepsilon(\mathbf{x})=\varepsilon_{\mathbf{x}},\quad \varepsilon_{\mathbf{x}}(f)=f(\mathbf{x}),$$

确定, 其中, $\mathbf{x}\in V,\quad f\in V^*,\varepsilon_{\mathbf{x}}\in V^{**}$.

证明 可直接验证ε的线性性. 事实上, 对所有的线性函数$f:V\to\mathfrak{K}$, 应有

$$\varepsilon_{\alpha\mathbf{x}+\beta\mathbf{y}}(f)=f(\alpha\mathbf{x}+\beta\mathbf{y})=\alpha f(\mathbf{x})+\beta f(\mathbf{y})=\alpha\varepsilon_{\mathbf{x}}(f)+\beta\varepsilon_{\mathbf{y}}(f)=(\alpha\varepsilon_{\mathbf{x}}+\beta\varepsilon_{\mathbf{y}})(f)$$

从而, $\varepsilon_{\alpha\mathbf{x}+\beta\mathbf{y}}=\alpha\varepsilon_{\mathbf{x}}+\beta\varepsilon_{\mathbf{y}}$, 也就是$\varepsilon(\alpha\mathbf{x}+\beta\mathbf{y})=\alpha\varepsilon(\mathbf{x})+\beta\varepsilon(\mathbf{y})$.

为了证明ε是个双射, 在V和V^*中选取对偶的基底$V=\langle \mathbf{e}_1,\cdots,\mathbf{e}_n\rangle$, $V^*=\langle e^1,\cdots,e^n\rangle$, 那么

$$\varepsilon_{\mathbf{e}_j}(e^i)=e^i(\mathbf{e}_j)=\delta_{ij}.$$

借助定理1的证明, 我们看到等式$V^{**}=\langle \varepsilon_{\mathbf{e}_1},\varepsilon_{\mathbf{e}_2},\cdots,\varepsilon_{\mathbf{e}_n}\rangle$, 也就是说, $(\varepsilon_{\mathbf{e}_j})$ 是V^{**}的对偶于(e^i)的基底. ε的单性和满性是显然的, 同构ε是自然的由其定义即可得出. □

定义4 *由V和V^{**}之间存在自然同构而显示出来的空间性质被称为***自反性**.

自反性使得V和V^{**}完全平等. 用定理2中的自然同构ε把V^{**}和V等同起来, 就可以把V看成是V^*的线性函数空间, 进而赋予联合公式(3)以新的意义: $\mathbf{x}(f)=(f,\mathbf{x})=f(\mathbf{x})$. 特别地, **对于$V^*$的所有的基底都必有唯一确定的与之对偶的$V$的基底.**

4. 线性无关性的判别法 利用对偶空间V^*的概念可以方便简洁地表达出空间V中向量无关性的各种判别方法.

先有

引理1 *如果$\mathbf{a}_1,\cdots,\mathbf{a}_m$是$V$中的线性相关的向量, 而$f_1,\cdots,f_m$是$V$上的任意的线性函数, 那么,*

$$\det(f_i(\mathbf{a}_j))=0,\quad 1\leqslant i,j\leqslant m$$

(i是行指标, j是列指标).

证明 由于$\mathbf{a}_1,\cdots,\mathbf{a}_m$的线性相关性, 其中必有一个, 比如说是$\mathbf{a}_m$; 必为其余向量的线性组合(§2之定理1), 设$\mathbf{a}_m=\alpha_1\mathbf{a}_1+\cdots+\alpha_{m-1}\mathbf{a}_{m-1}$, 在行列式$\det(f_i(\mathbf{a}_j))$中, 从最后一列减去第1列乘$\alpha_1$, 第2列乘$\alpha_2,\cdots$, 最后, 用$\alpha_{m-1}$乘第$m-1$列, 我们知道, 经过这些变换行列式值不变, 我们计算出, 最后一列的第i个位置上就是$f_i(\mathbf{a}_m)-\alpha_1 f_i(\mathbf{a}_1)-\cdots-\alpha_{m-1}f_i(\mathbf{a}_{m-1})=f_i(\mathbf{a}_m-\alpha_1\mathbf{a}_1-\cdots-\alpha_{m-1}\mathbf{a}_{m-1})=f_i(\mathbf{0})=0, i=1,2\cdots,m$. 所以, 行列式等于零. □

引理2 *如果$(f_1,\cdots,f_n)$是向量空间V的对偶空间V^*的一个基底, 那么, 向量$\mathbf{a}_1,\cdots,\mathbf{a}_n\in V$线性无关当且仅当*

$$\det(f_i(\mathbf{a}_j))\neq 0.\quad 1\leqslant i,j\leqslant n.$$

证明 由引理1, $\mathbf{a}_1,\cdots,\mathbf{a}_n$线性相关导致行列式的等式为零. 现在设它们线性无关, 即$V=\langle \mathbf{a}_1,\cdots,\mathbf{a}_n\rangle$. 用$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$代表$V$的对偶于$(f_1,\cdots,f_n)$的基底, 而用$\alpha_{1j},\cdots,\alpha_{nj}$代表向量$\mathbf{a}_j$在这个基底之下的坐标. 那么

$$\begin{pmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn}\end{pmatrix}$$

就是由基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$到基底$(\mathbf{a}_1,\cdots,\mathbf{a}_n)$的转换矩阵. 据§2之定理4, 它是可逆的, 从而$\det(\alpha_{ij})\neq 0$. 但$\alpha_{ij}=f_i(\mathbf{a}_j)$(见(4)), 继而得$\det(f_i(\mathbf{a}_j))=0$. □

定理3 设$(f_1, \cdots, f_n)$是对偶于V的空间V^*的一个基底. 那么, 向量组$\mathbf{a}_1, \cdots, \mathbf{a}_k \in V$ 的秩等于所有形如

$$\det(f_i(\mathbf{a}_j)), \quad 1 \leqslant i = i_1, \cdots, i_m \leqslant n; 1 \leqslant j = j_1, \cdots, j_m \leqslant k. \tag{5}$$

的非零行列式的最大阶数.

证明 用r代表向量组$\mathbf{a}_1, \cdots, \mathbf{a}_k$的秩.对任意$m > r$, $\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_m}$必线性相关.据引理1, 这意味着, 阶数$m > r$的形如(5)的行列式必然等于零.

剩下的是要证明, 存在一个形如(5)的非零行列式, 其秩为r.为此, 用$\overline{f_1}, \cdots, \overline{f_n}$代表线性函数$f_1, \cdots, f_n$在子空间$U = \langle \mathbf{a}_1, \cdots, \mathbf{a}_k \rangle$上的限制. 先证明

$$\langle \overline{f_1}, \cdots, \overline{f}_n \rangle = U^*, \tag{6}$$

其中U^*是对偶于U的子空间.

事实上, 显然$\langle \overline{f_1}, \cdots, \overline{f_n} \rangle \subseteq U^*$. 其次, 设$\widetilde{f}$是$U^*$的任意一个向量, $(\mathbf{e}_1, \cdots, \mathbf{e}_r)$是$U$的一个基底, 而$(\mathbf{e}_1, \cdots, \mathbf{e}_r; \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n)$是它在$V$中的扩展基底. 看线性函数$f \in V^*$, 使得$f(\mathbf{e}_i) = \widetilde{f}(\mathbf{e}_i)$, $i = 1, \cdots, r$; $f(\mathbf{e}_i) = 0, i = r+1, \cdots, n$ (在基底向量处取任意值的线性函数$f \in V^*$都存在, 从而有的可取成现在这个样). 因为$V^* = \langle f_1, \cdots, f_n \rangle$, 故$f = \beta_1 f_1 + \cdots + \beta_n f_n$. 把这个等式中的所有函数都限制在$U$上. 显然, $\overline{f} = f|_U = \widetilde{f}$. 因为$\overline{f}$和$\widetilde{f}$在空间$U$的基底向量$\mathbf{e}_1, \cdots, \mathbf{e}_r$上都取相同的值, 于是$\widetilde{f} = \overline{f} = \beta_1 \overline{f}_1 + \cdots + \beta_n \overline{f}_n$, 随之就有$\widetilde{f} \in \langle \overline{f_1}, \cdots, \overline{f_n} \rangle$, 也就是$U^* \subseteq \langle \overline{f_1}, \cdots, \overline{f_n} \rangle$, 从而证明了(6).

最后, 在$\mathbf{a}_1, \cdots, \mathbf{a}_k$中选择$r$个线性无关的向量(设为$\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_r}$), 在$\overline{f}_1, \cdots, \overline{f}_n$中选择$r$个线性无关的向量(设为$\overline{f}_{i_1}, \cdots, \overline{f}_{i_r}$), 它们分别组成子空间$U$和$U^*$的基底, 据引理2

$$\det(\overline{f}_i(\mathbf{a}_j)) \neq 0, \quad i = i_1, \cdots, i_r; \quad j = j_1, \cdots, j_r.$$

余下的事情是只需注意$\overline{f}_i(\mathbf{a}_j) = f_i(\mathbf{a}_j)$. □

我们又一次接触矩阵秩的概念(见[BA I]第2章§2), 但再次停留在其性质上则没有意义.

5. 齐次线性方程组解的几何解释 回想一下, n个未知量的齐次线性方程组相容时, 把它的解可解释成基础域$\mathfrak{K}$上列(或行)空间$\mathfrak{K}^n$的向量(我们已经这样做了), 那么, 在$\mathfrak{K}^n$中由这个方程组的解生成一个子空间.采用更加抽象的观点, 据$m \times n$的齐次线性方程组的定义, 可写成

$$f_1(\mathbf{x}) = 0, \cdots, f_m(\mathbf{x}) = 0, \tag{7}$$

其中$\mathbf{x}$是n维空间V的向量, 而$f_1, \cdots, f_m \in V^*$.为了返回到通常的记法, 只要在V中任意选取一个基底.

定理4 i) 如果向量组$f_1, \cdots, f_m \in V^*$的秩为r, 那么, 齐次线性方程组(7)的解的子空间$U \subset V$的维数等于$n - r$ $(n = \dim_{\mathfrak{K}} V)$.

ii) 任意子空间$U\subset V$必然是某个组(6)的解子空间.

证明 结论i), 在[BA I]第2章§3中已经证明过了, 而现在的推断领会起来会更自然.去掉一般性的约束, 假定$f_1,\cdots,f_r$线性无关.那么, 余下的f_i都是它们的线性组合, 但方程组(7)实际上等价于

$$f_1(\mathbf{x})=0,\cdots,f_r(\mathbf{x})=0. \tag{7'}$$

把$f_1,\cdots,f_r$扩展成空间V^*的基底(当$i\leqslant r$时, $e^i=f_i$)$(e^1,\cdots,e^n)$.设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是$V$的对偶于$(e^1,\cdots,e^n)$的基底.那么, 方程组(7′)中, 任意$\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n$必然有$x_1=\cdots=x_r=0$.从而, 方程组(7′)的解空间$U$由形如$\mathbf{x}=x_{r+1}\mathbf{e}_{r+1}+\cdots+x_n\mathbf{e}_n$的向量组成, 即$U=\langle\mathbf{e}_{r+1},\cdots,\mathbf{e}_n\rangle$. 注意, $x_{r+1},\cdots,x_n$在起无关的自由变元的作用. 因为$\mathbf{e}_{r+1},\cdots,\mathbf{e}_n$是线性无关的, 所以, $\dim U=n-r$.

ii) 设$(\mathbf{e}_1,\cdots,\mathbf{e}_s)$是子空间$U\subset V$的一个基底, 也是整个空间$V$的一个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$的一部分. 向量$\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n$属于$U$, 当且仅当, $x_{s+1}=\cdots=x_n=0$. 选取V^*的一个与$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$对偶的基底$(f_1,\cdots,f_n)$. 那么, $x_i=f_i(\mathbf{x})$, 进而条件$\mathbf{x}\in U$可写成$f_{s+1}(\mathbf{x})=0,\cdots,f_n(\mathbf{x})=0$. □

习　题

1. 正如由定义就很容易算出来(见§1之习题1)一样, 迹函数tr: $X\mapsto \mathrm{tr}\,X$在域$\mathfrak{K}$上所有n阶方阵的空间$V=M_n(\mathfrak{K})$上是线性的, 证明, V上的每个线性函数f必形如$f(X)=\mathrm{tr}AX$, 其中矩阵$A=A_f$是唯一确定的.

2. 设$a(t)$是$\mathbb{R}[t]$中的一个固定的多项式, P_n是所有次数$\leqslant n-1$的实的多项式子空间, 在P_n上研究下列函数:

(1) $f(u)=\int_{-1}^1 a(t)u(t)dt;\quad u(t)\in P_n;$

(2) $f(u)=\int_0^1 a(t)u(t^2)dt;$

(3) $f(u)=\int_0^1 a(t)[u(t)]^2dt;$

(4) $f(u)=\dfrac{d^3}{dt^3}a(t)u(t)|_{t=-1}.$

3. 设V是向量空间, 再设$f,\ g\in V^*$且Ker f=Ker g.证明, 必有某个纯量λ使得$g=\lambda f$.

4. 设$\mathbf{x}$是向量空间V的一个非零向量. 条件$f(\mathbf{x})$=1能将函数$f\in V^*$唯一决定吗?

5. 证明, 在域$\mathfrak{K}$上的n维空间V中, 对任意非零的线性函数f都能找到空间V的一个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$使得

$$f(\alpha_1\mathbf{e}_1+\cdots+\alpha_n\mathbf{e}_n)=\alpha_i,\quad \forall\alpha_i\in\mathfrak{K}.$$

§4 双线性型和二次型

1. 多重线性映射 初次阅读可将本节省略.已经显露出其功能的余向量(V上的线性函数)概念还可以用到更一般情形.

研究$\mathfrak{K}$上向量空间$V_1,\cdots,V_p;U$. 映射

$$f:V_1\times V_2\times\cdots\times V_p\to U$$

称为是**多重线性的**(在已给定的情形, p-**线性**), 如果对任意指数$i=1,\cdots,p$及任意固定的向量$\mathbf{a}_j\in V_j, 1\leqslant j\leqslant p, j\neq i$, 映射

$$f_i:\mathbf{v}\mapsto f(\mathbf{a}_1,\cdots,\mathbf{a}_{i-1};\mathbf{v},\mathbf{a}_{i+1},\cdots,\mathbf{a}_p)$$

都是线性型(线性函数), 即

$$f_i(\alpha\mathbf{x}+\beta\mathbf{y})=\alpha f_i(\mathbf{x})+\beta f_i(\mathbf{y}),\quad \forall\mathbf{x},\mathbf{y}\in V_i,\quad \alpha,\beta\in\mathfrak{K}. \tag{1}$$

在第2章, 我们会继续研究(1)型线性映射.现在仅仅得到一个评注. 与线性函数一样, 容易确信, 两个p-线性映射的线性组合$\alpha f+\beta g$也是个p-线性映射. 这个情况把所有$V_1\times\cdots\times V_p\to U$的$p$-线性映射的集合$\mathcal{L}(V_1,\cdots,V_p;U)$作为$\mathfrak{K}$上的线性空间加以研究.

取$V_1=\cdots=V_p=U=\mathfrak{K}$ (1维向量空间), 我们就得到一个最简单的例子, 令

$$f(v_1,\cdots,v_p)=v_1\cdots v_p$$

即可. 更一般地, 任意$V_1\times\cdots\times V_p$到$\mathfrak{K}$的多重线性映射都被称为$V_1\times\cdots\times V_p$上的**多重线性型**. 如果说$l^i$: $\mathbf{v}_i\mapsto l^i(\mathbf{v}_i)$, $i=1,\cdots,p$是V_i上的线性函数, 那么, 用关系式

$$f(\mathbf{v}_1,\cdots,\mathbf{v}_p)=l^1(\mathbf{v}_1)\cdots l^p(\mathbf{v}_p),$$

定义的函数f就是$V_1\times\cdots\times V_p$上的多重线性函数. 称它为线性函数(型)$l^1,\cdots,l^p$的**张量积**且表成$f=l^1\otimes l^2\otimes\cdots\otimes l^p$或简记为$l^1l^2\cdots l^p$ (有序).

容易证明, $V_1\times\cdots\times V_p$上任意的一个多重线性型必为线性函数的张量积. 但眼下我们不需要这个事实. 在$V_1=\cdots=V_p=V$时, $V^p=V\times\cdots\times V$ (集合V的p个元素的笛卡儿积), 此时, 记

$$\mathcal{L}_p(V,\mathfrak{K})=\mathcal{L}(V,\cdots,V;\mathfrak{K})$$

是很方便的.

在$V^p\times V^{*q}$上的多重线性型随后将被说成是$p+q$价的(p,q)型张量. 于是(0, 1)型张量就可以看成是V中的向量, 这是§3的第3目给出的自反性.

V^p上的多重线性型f($(p,0)$型张量)称为**对称的**. 如果对任意$\mathbf{v}_1,\cdots,\mathbf{v}_p\in V$及任意置换$\pi\in S_p$, 都有

$$f(\mathbf{v}_{\pi(1)},\mathbf{v}_{\pi(2)},\cdots,\mathbf{v}_{\pi(p)})=f(\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p).$$

如果

$$f(\mathbf{v}_{\pi(1)},\mathbf{v}_{\pi(2)},\cdots,\mathbf{v}_{\pi(p)})=\varepsilon_\pi f(\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p),$$

其中ε_π是置换的奇偶性, 那么, 称f为**斜对称的**(或**交错的**)线性型. 在$\dim V=p$的情形, 我们知道一个很好的例子, 就是将矩阵的行列式看成列或行的函数. 在这里给出

一般性定义的目的仅仅是把所有我们已经遇到过的和要进一步研究的局部性的概念放到一般性的框框里. 一般情形的张量记号我们只在第6章里应用.

2. 双线性型 现在, 限制在$V_1 = V_2 = V$的情形, 而且说f是V上的双线性型(p=2)(而不是更准确地说V^2), 与通常定义中的关系式相对应, 域$\mathfrak{K}$上向量空间V上的**双线性型**f由下列性质来刻画, 即

$$\begin{aligned} f(\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{w}) &= \alpha f(\mathbf{u},\mathbf{w})+\beta f(\mathbf{v},\mathbf{w}), \\ f(\mathbf{w},\alpha\mathbf{u}+\beta\mathbf{v}) &= \alpha f(\mathbf{w},\mathbf{u})+\beta f(\mathbf{w},\mathbf{v}). \end{aligned} \tag{2}$$

对所有$\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}\in V, \alpha, \beta\in\mathfrak{K}$都成立. 注意, 一般说来, $f(\mathbf{u}, \mathbf{v})\neq f(\mathbf{v}, \mathbf{u})$.

在V中选择某个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$, 并通过其坐标,

$$\mathbf{x} = x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n, \qquad \mathbf{y} = y_1\mathbf{e}_1+\cdots+y_n\mathbf{e}_n,$$

选出向量$\mathbf{x}$, $\mathbf{y}\in V$. 我们利用定义的性质(2), 用n^2个纯量$f(\mathbf{e}_i,\mathbf{e}_j)$ 去记线性型$f(\mathbf{u}, \mathbf{v})$.即

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &= f(\sum_i x_i\mathbf{e},\sum_j y_j\mathbf{e}_j) = \sum_i x_i f(\mathbf{e}_i,\sum_j y_j\mathbf{e}_j) \\ &= \sum_i x_i\sum_j y_j f(\mathbf{e}_i,\mathbf{e}_j) = \sum_{i,j} f_{ij}x_iy_j, \end{aligned} \tag{3}$$

其中, $f_{ij} = f(\mathbf{e}_i,\mathbf{e}_j)$.

矩阵$F = (f_{ij})$称为V上双线性型f在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下的矩阵. 引入对于$n\times 1$阶坐标矩阵(列)$X$=$[x_1,x_2,\cdots,x_n]$和它的转置$1\times n$阶矩阵${}^tX = (x_1,x_2,\cdots,x_n)$ (行), 我们就可以把表达式(3)写成

$$f(\mathbf{x},\mathbf{y}) = {}^tX\cdot F\cdot Y \tag{4}$$

的形式. 为此只需利用$1\times n, n\times n, n\times 1$阶矩阵的已知的乘法规则.

反之, 如果有方阵$F = (f_{ij})$, 借助关系式(4) (或者(3)), 令$f(\mathbf{e}_i,\mathbf{e}_j) = f_{ij}$, 即可定义$V$上一个双线性型$f$. 因此, 在给定的$\mathfrak{K}$上向量空间的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下, 有一个$\mathfrak{K}$上$n\times n$矩阵和$V(\dim_{\mathfrak{K}}V = n)$上双线性型之间的相互单值的对应. 事实上, 这个对应是所有$V\times V\to\mathfrak{K}$双线性型的向量空间$\mathcal{L}_2(V,\mathfrak{K})$ (如果f, $g\in\mathcal{L}_2(V,\mathfrak{K})$, 那么$\alpha f+\beta g\in\mathcal{L}_2(V,\mathfrak{K})$, 这是显而易见的)到$\mathfrak{K}$上所有$n$阶矩阵的向量空间$M_n(\mathfrak{K})$的同构. 实际上, 如果

$$f(\mathbf{x},\mathbf{y}) = {}^tXFY, \quad g(\mathbf{x},\mathbf{y}) = {}^tXGY,$$

那么,

$$\alpha f(\mathbf{x},\mathbf{y})+\beta g(\mathbf{x},\mathbf{y}) = {}^tX\cdot(\alpha F+\beta G)\cdot Y.$$

3. 双线性型的矩阵的转换规则 双线性型f的公理性定义与V中怎样选择基底没有关系. 为了使f 的矩阵记法有真实价值, 需要补充对应$f\mapsto F$ 在向新的基底转换

时矩阵F 的转换规则. 设与V 的给定的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 并列的还有一个基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$, 并有转换矩阵$A=(a_{ij})$:

$$\mathbf{e}'_j=\sum_{i=1}^{n}a_{ij}\mathbf{e}_i,\qquad j=1,\cdots,n.$$

如果$x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n=\mathbf{x}=x'_1\mathbf{e}'_1+\cdots+x'_n\mathbf{e}'_n$, 那么, 坐标列$X$和$X'$可用关系式$X=AX'$联系起来. 现在, 设$F=(f_{ij})$是双线性型$f$在基底$(\mathbf{e}_i)$之下的矩阵, 而$F'=(f'_{ij})$是同一线性型在基底$(\mathbf{e}'_i)$之下的矩阵, 也就是说, $f_{ij}=f(\mathbf{e}_i,\mathbf{e}_j)$和$f'_{ij}=f(\mathbf{e}'_i,\mathbf{e}'_j)$. 因为, ${}^t(AX')={}^tX'\cdot{}^tA$, 又因为值$f(\mathbf{x},\mathbf{y})$一般地与基底的选择没有关系, 所以,

$$\begin{aligned}{}^tX'\cdot F'\cdot Y'&=f(\mathbf{x},\mathbf{y})={}^tX\cdot F\cdot Y\\&={}^t(AX')\cdot F\cdot(AY')={}^tX'\cdot{}^tA\cdot F\cdot A\cdot Y'.\end{aligned}$$

比较这个等式的左右两端, 我们即得到结论, 有

定理1 V上双线性型f在基底$(\mathbf{e}_i)$和$(\mathbf{e}'_i)$之下的矩阵F和F'可由关系式

$$F'={}^tA\cdot F\cdot A \tag{5}$$

联系起来, 其中A是从$(\mathbf{e}_i)$到$(\mathbf{e}'_i)$的**转换矩阵**.

定义1 矩阵F和$F'={}^tAFA$, 而$\det A\neq 0$即说成是**合同的**, 双线性型f在任意一个基底$(\mathbf{e}_i)$之下对应的矩阵F的秩即称为f的**秩**.

推论 双线性型f的秩rank f是它的一个不变量, 与基底的选择无关.

证明 把[BA Ⅰ]第2章§3之定理5的推论1用到合同矩阵(5)上. □

关于双线性型的秩的论断还有一个. 用L_f代表所有那样的$\mathbf{x}\in V$的集合, 它对所有$\mathbf{y}\in V$都有$f(\mathbf{x},\mathbf{y})=0$, 即$f(\mathbf{x},V)=0$. 简易的检验表明, L_f是V的子空间. 称它是f的**左根**或者f的**左核**. 显然, $\dim L_f$是个只依赖于f的量. 设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是$V$的一个基底. 条件$\mathbf{x}\in L_f$等价于

$$f(\mathbf{x},\mathbf{e}_1)=0,\cdots,f(\mathbf{x},\mathbf{e}_n)=0.$$

这是一个由线性函数$\mathbf{x}\mapsto f_j(\mathbf{x})=f(\mathbf{x},\mathbf{e}_j)=0,j=1,\cdots,n.$ 决定的一个方程组.

函数f_j的坐标是纯量$f_j(\mathbf{e}_i)$, 即系数$f(\mathbf{e}_i,\mathbf{e}_j)=f_{ij}$, 也就是$F$的第$j$行. 因为线性函数组$f_1,\cdots,f_n\in V^*$的秩与矩阵$F=(f_{ij})$的秩一样, 所以, 据[BA Ⅰ]第2章§3的定理7, 有等式$\dim L_f=n-r$. 换言之,

$$r=\dim V-\dim L_f$$

是个不依赖于任何基底的量.

4. 对称型与斜对称型 与第1目相对应地, 双线性型$f:V\times V\to\mathcal{K}$称为**对称的**, 如果对所有$\mathbf{x},\mathbf{y}\in V$都有$f(\mathbf{x},\mathbf{y})=f(\mathbf{y},\mathbf{x})$; 称为**斜对称的**, 如果$f(\mathbf{x},\mathbf{y})=-f(\mathbf{y},\mathbf{x})$. 这

些术语与对称多项式和斜对称多项式的概念很是一致(见[BA I]), 也同样与**对称矩阵**$A=(a_{ij})$, ${}^tA=A$, **斜对称矩阵**${}^tA=-A$的概念一致. 因为$f(\mathbf{y},\mathbf{x})={}^tf(\mathbf{y},\mathbf{x})$($1\times1$阶转置, 纯量), 故由$f(\mathbf{x},\mathbf{y})=\varepsilon f(\mathbf{y},\mathbf{x}),\varepsilon=\pm1$, 与关系(4)对应, 有

$$\begin{aligned}{}^tX\cdot F\cdot Y&=f(\mathbf{x},\mathbf{y})=\varepsilon f(\mathbf{y},\mathbf{x})=\varepsilon\cdot{}^tf(\mathbf{y},\mathbf{x})\\&=\varepsilon\cdot{}^t({}^tY\cdot F\cdot X)\\&=\varepsilon\cdot{}^tX\cdot{}^tF\cdot Y,\end{aligned}$$

从而${}^tF=\varepsilon F$. 反之, 如果${}^tF=\varepsilon F,\varepsilon=\pm1$, 那么, 矩阵$F$对应的双线性型$f$必然满足关系式$f(\mathbf{x},\mathbf{y})=\varepsilon f(\mathbf{y},\mathbf{x})$.

剩下来要补充的是, 按照(5)

$$F'={}^tA\cdot F\cdot A\Rightarrow{}^tF'={}^tA\cdot{}^tF\cdot A=\varepsilon\cdot{}^tA\cdot F\cdot A=\varepsilon F',$$

所以, 对于f的矩阵F的对称性和斜对称性与基底选择无关. 因此, 双线性型f是对称的或者是斜对称的, 只要它在V的任意一个基底之下是对称的或者是斜对称的.

定理2 如果char $\mathfrak{K}\neq2$, 那么, 所有双线性型构成的空间$\mathcal{L}_2(V,\mathfrak{K})$是对称型子空间$\mathcal{L}_2^+(V,\mathfrak{K})$和斜对称型空间$\mathcal{L}_2^-(V,\mathfrak{K})$的直和.

$$\mathcal{L}_2(V,\mathfrak{K})=\mathcal{L}_2^+(V,\mathfrak{K})\oplus\mathcal{L}_2^-(V,\mathfrak{K})$$

证明 如果$f\in\mathcal{L}_2^+(V,\mathfrak{K})\cap\mathcal{L}_2^-(V,\mathfrak{K})$, 那么

$$f(\mathbf{x},\mathbf{y})=f(\mathbf{y},\mathbf{x})=-f(\mathbf{x},\mathbf{y})\Rightarrow2f(\mathbf{x},\mathbf{y})=0\Rightarrow f(\mathbf{x},\mathbf{y})=0$$

(因为有条件char $\mathfrak{K}\neq2$), 从而$f=0$. 所以, 和$\mathcal{L}_2^++\mathcal{L}_2^-$是直和.

另一方面, 关系式

$$f(\mathbf{x},\mathbf{y})=\frac{1}{2}\{f(\mathbf{x},\mathbf{y})+f(\mathbf{y},\mathbf{x})\}+\frac{1}{2}\{f(\mathbf{x},\mathbf{y})-f(\mathbf{y},\mathbf{x})\}$$

或者相应的矩阵关系式

$$F=\frac{1}{2}(F+{}^tF)+\frac{1}{2}(F-{}^tF)$$

都表明所有的双线性型f都可以表示成一个对称型和一个斜对称型之和的形式. □

在域$\mathfrak{K}=Z_2$上, 每个斜对称矩阵同时也是对称的, 从而定理的论断不再真确. 因为, 举个例子来说吧, 矩阵$\begin{pmatrix}0&1\\0&1\end{pmatrix}$就不是对称的. 但仍有交错双线性型$f$的概念, $f(\mathbf{u},\mathbf{u})=0,\forall\mathbf{u}\in V$. 但, 它在char $\mathfrak{K}\neq2$的情形与斜对称型的概念一致(验证它). 今后, 我们总假定char $\mathfrak{K}\neq2$.

5. 二次型 考察对称的双线性型引出的下列重要概念, 它是在很多数学分支中以极自然的方式产生的.

定义2 域$\mathcal{K}$上有限维空间V上的函数$q:V\to\mathcal{K}$, 它如果具有如下两个性质:

i) $q(-\mathbf{v})=q(\mathbf{v}),\quad \forall \mathbf{v}\in V$;

ii) 由公式

$$f(\mathbf{x},\mathbf{y})=\frac{1}{2}\{q(\mathbf{x}+\mathbf{y})-q(\mathbf{x})-q(\mathbf{y})\}, \tag{6}$$

决定的映射$f:V\times V\to\mathcal{K}$是$V$上的双线性型(显然是对称的), 则称$q$是$V$上的**二次型**, 同样地, 称$f$的秩为$q$的秩: rank q=rank f.

还要说明, 利用公式(6), 由q得到的对称的双线性型f是**极化的**; 或者, f是与二次型q**配极**的双线性型.

现在设f是V上任意一个对称的双线性型, 令

$$q_f(\mathbf{x})=f(\mathbf{x},\mathbf{x}), \tag{7}$$

就得到了一个满足二次型定义中条件i), ii)的函数$q_f:V\to\mathcal{K}$, 因为$f(-\mathbf{x},-\mathbf{x})=f(\mathbf{x},\mathbf{x})$而且

$$f(\mathbf{x},\mathbf{y})=\frac{1}{2}\{f(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})-f(\mathbf{x},\mathbf{x})-f(\mathbf{y},\mathbf{y})\}. \tag{8}$$

可能会想到, q_f是个如此特别的二次型. 实际上并非如此, 因为下面的定理成立(可以省略它的并不复杂的证明而不影响此后的理解).

定理3 每个二次型q都可以按着自己的配极双线性型f唯一地恢复原型; 换言之, q=q_f.

证明 在(6)中令$\mathbf{y}=-\mathbf{x}$:

$$-f(\mathbf{x},\mathbf{x})=\frac{1}{2}\{q(\mathbf{0})-q(\mathbf{x})-q(-\mathbf{x})\},$$

从而

$$q(\mathbf{x})=f(\mathbf{x},\mathbf{x})+\frac{1}{2}q(\mathbf{0}).$$

因为f是个双线性型, 所以, $f(\mathbf{0},\mathbf{0})=0$.因为, 当$\mathbf{x}=\mathbf{0}$时有$q(\mathbf{0})=\frac{1}{2}q(\mathbf{0})$, 也就是$q(\mathbf{0})=0$, 这就是说, $q(\mathbf{x})=f(\mathbf{x},\mathbf{x})$. □

定义3 称与q配极的双线性型f在空间V的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下的矩阵$F$是二次型$q$=$q_f$的矩阵.

可见, $F=(f_{ij})$, 其中

$$f_{ij}=\frac{1}{2}\{q(\mathbf{e}_i+\mathbf{e}_j)-q(\mathbf{e}_i)-q(\mathbf{e}_j)\}\quad i,j=1,2,\cdots,n.$$

任意一个对称矩阵F=(f_{ij})自己也适应由关系式

$$q(\mathbf{x})={}^tX\cdot F\cdot X=\sum_{i,j}f_{ij}x_ix_j \tag{9}$$

给定的二次型q.

这样一来, 与二次型名称相适应的是向量$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$坐标的齐次二次函数. 要突出关注对角矩阵.

定义4 称二次型q在V的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下具有**规范型**或**对角型**, 如果对任意向量$\mathbf{x} = \sum x_i\mathbf{e}_i \in V$, $q(\mathbf{x})$的值可用公式

$$q(\mathbf{x}) = \sum_i f_{ii} x_i^2$$

计算. 此时, 基底$(\mathbf{e}_i)$称为对q而言的**规范基底**.

这个术语也和对应的与之配极的双线性型f有关系

$$f(\mathbf{x}, \mathbf{y}) = \sum_i f_{ii} x_i y_i.$$

并不要求二次型的规范型或者它的规范基底唯一确定. 比方说, 对任意规范基底中的向量施以任意的置换仍然得规范基底.

注意, 在规范型中, rank q_f=rank f简单地就是非零的系数f_{ii}的个数. 同时, 对照第3目结尾处的说明, rank q = dim $V-$ dim L_q, 其中$L_q = L_f$是型f的核(或根)(由于f是对称的, 左右根无区别). 子空间$L_q \subset V$同样地被称为是二次型q的**迷向**(或**零**)**子空间**, 用q的术语, 如下方式定义

$$L_q = \{\mathbf{u} \in V | q(\mathbf{u} + \mathbf{v}) = q(\mathbf{u}) + q(\mathbf{v}),\ \forall \mathbf{v} \in V\}.$$

二次型q的秩是个不变量.

6. 二次型的规范型 关于选取基底使得给定的一个二次型在此基之下有形式比较简单, 所谓规范形式的问题在理论上和应用中都有重要价值.

定理4 对V上每个对称的双线性型f都有规范基底.

证明 当n=1时, 命题是显然的, 所以, 可以对n用数学归纳法. 如果$f(\mathbf{x}, \mathbf{y})$=0, 对所有的$\mathbf{x}, \mathbf{y} \in V$都成立(即$f$=0), 那么, 定理显然: 任意一个基底都适用. 如果$f \neq 0$, 那么, 对应的二次型也不为零(等式(6), (8)或定理3). 设$\mathbf{e}_1$是这样一个向量, $f(\mathbf{e}_1, \mathbf{e}_1) = q(\mathbf{e}_1) \neq 0$. 于是, 线性函数$f_1 : \mathbf{x} \mapsto f(\mathbf{x}, \mathbf{e}_1)$非零$(f_1(\mathbf{e}_1) \neq 0)$, 据§3定理4, 线性子空间

$$L = \mathrm{Ker} f_1 = \{\mathbf{x} \in V | f_1(\mathbf{x}) = 0\}$$

的维数是$n-1$, 它是个超平面, 据归纳法假设对L必有基底$(\mathbf{e}_2, \cdots, \mathbf{e}_n)$, 在此基之下, f限制在L上的矩阵是对角的, 即

$$f(\mathbf{e}_i, \mathbf{e}_j) = 0, \qquad i, j = 2, \cdots, n, \qquad i \neq j.$$

按构作方式, $f(\mathbf{e}_i, \mathbf{e}_1) = 0$, $i = 2, 3, \cdots, n$. 所以得到性质$f(\mathbf{e}_i, \mathbf{e}_j) = 0$, $i \neq j$. 所以只要向量组$\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$线性无关, $(\mathbf{e}_k)$就具有规范基底的特征了. 设, 情形相反, 那么, 在任意非平凡关系式

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0}$$

中, 只能有$\alpha_1 \neq 0$, 因为$(\mathbf{e}_2, \cdots, \mathbf{e}_n)$是$L$的基底, 但这种情况$\mathbf{e}_1 = \sum\limits_{i>1} \beta_i \mathbf{e}_i$且

$$0 \neq f_1(\mathbf{e}_1) = f_1\left(\sum_{i>1} \beta_i \mathbf{e}_i\right) = \sum_{i>1} \beta_i f_1(\mathbf{e}_i) = 0,$$

这是一个矛盾, 定理得证. □

推论1　在域$\mathbb{K}$上的n维向量空间V上给定一个秩为$r \leqslant n$的二次型q. 那么, 在V中存在一个基底$(\mathbf{e}_i)$, 在它之下q取规范形式:

$$q(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_r x_r^2. \tag{10}$$

推论2　对每个对称矩阵F都有非退化矩阵A使得${}^tAFA$是和F秩相同的对角矩阵, 换言之, 每个对称矩阵都合同于某个对角矩阵.

前面见到过的把双线性型(因而二次型)化成规范型的归纳方法是拉格朗日(1736—1813)提出来的. 自然地, 在实际应用中, 采取按某个顺序作用的坐标记法. 由可解释成n个无关变量的齐次二次多项式的二次型$q(\mathbf{x})$的表达式(9)

$$q(x_1, \cdots, x_n) := q(\mathbf{x}) = \sum_{i,j=1}^{n} f_{ij} x_i x_j,$$

出发, 用古巴比伦补充方法, 摆脱混合项$x_i x_j, i \neq j$而转化成完全二次项, 提出所有含x_1的项:

$$q(x_1, \cdots, x_n) = f_{11} x_1^2 + 2f_{12} x_1 x_2 + 2f_{13} x_1 x_3 + \cdots + 2f_{1n} x_1 x_n + \sum_{i,j\neq 1} f_{ij} x_i x_j$$

(易见, 有形如$f_{1j} x_1 x_j + f_{j1} x_j x_1$的和, 而$f_{j1} = f_{1j}$, 所以出现2倍积). 开始时, 设$f_{11} \neq 0$, 且依靠不含$x_1$的项的系数提出来完全项

$$q(x_1, \cdots, x_n) = \frac{1}{f_{11}}(f_{11} x_1 + f_{12} x_2 + \cdots + f_{1n} x_n)^2 + \sum_{i,j\neq 1} f'_{ij} x_i x_j.$$

现在, 设

$$x'_1 = f_{11} x_1 + f_{12} x_2 + \cdots + f_{1n} x_n, \qquad x'_i = x_i \quad i > 1,$$

我们就把二次型q化成了

$$q(x_1, \cdots, x_n) = \frac{1}{f_{11}}(x'_1)^2 + q'(x'_2, \cdots, x'_n)$$

的形式. 其中$q'(x_2, \cdots, x_n) = \sum\limits_{i,j=2}^{n} f'_{ij} x'_i x'_j$是变量数目更少的一个二次型. 设$f'_{22} \neq 0$, 把它写成

$$\begin{aligned} q'(x'_2, \cdots, x'_n) &= f'_{22} x_2'^2 + f'_{23} x'_2 x'_3 + \cdots + f'_{2n} x'_2 x'_n + \sum_{i,j>2} f'_{ij} x'_i x'_j \\ &= \frac{1}{f'_{22}}(f'_{22} x'_2 + f'_{23} x'_3 + \cdots + f'_{2n} x'_n)^2 + \sum_{i,j>2} f''_{ij} x'_i x'_j \end{aligned}$$

的形式(以提出新的完全项为前提, 把f'_{ij}转变成f''_{ij}), 依次替换变量.

$$x''_1 = x'_1, \quad x''_2 = f'_{22}x'_2 + f'_{23}x'_3 + \cdots + f'_{2n}x'_n, \quad x''_i = x'_i, \quad i > 2,$$

就提供给我们一个表达式

$$q(\mathbf{x}) = \frac{1}{f_{11}}(x''_1)^2 + \frac{1}{f'_{22}}(x''_2)^2 + q''(x''_3, \cdots, x''_n),$$

其中, $q''(x''_3, \cdots, x''_n) = \sum_{i,j=3}^{n} f''_{ij}x''_i x''_j$是一个变量数目更加少些的二次型.

这个过程显然可以一直继续到把$q(\mathbf{x})$写成$r=\text{rank } q$个二次项的线性组合. 按事情进展产生的变量替换是非退化的并且适应向新的基底的转换. 再做一点注释, 首先看, 限制性假设$f_{11} \neq 0, f'_{22} \neq 0, \cdots$ 不成立的情形. 若$f_{11} = 0$而对某个k有$f_{kk} \neq 0$, 那么可变动x_1, x_k的指标(或者变动基底向量的顺序). 但, 如果$q(\mathbf{x}) \neq 0$, 但不含任何二次项, 即对所有的k都有$f_{kk} = 0$, 那么, 不失一般性, 设$2f_{12}x_1x_2 \neq 0$, 此时可用替换

$$x_1 = x'_1 + x'_2, \quad x_2 = x'_1 - x'_2, \quad x_k = x'_k, \quad k > 2.$$

就有不可能被约掉的项$2f_{12}\left(x'^2 - x'^2_2\right)$出现了, 它提供启动我们的过程的可能性.

7. 实二次型 在任意域$\mathfrak{K}$ (限制char $\mathfrak{K} \neq 2$)上操作, 一般说来, 我们不可能把对角的二次型归结成更简单的形式. 但, 如果$\mathfrak{K} = \mathbb{R}$, 就可以让(10)中的所有系数均为 ±1. 实际上, 适当地置换基底的向量, 我们有权认为公式(10)的前s个系数$\lambda_1, \cdots, \lambda_s$ 是正的, 而其余的系数是负的. 在坐标替换

$$x'_i = \sqrt{\lambda_i} \cdot x_i, \quad 1 \leqslant i \leqslant s; \quad x'_i = \sqrt{-\lambda_i} \cdot x_i, \quad s+1 \leqslant i \leqslant r;$$
$$x'_i = x_i \quad r+1 \leqslant i \leqslant n,$$

之下即得到

$$q(\mathbf{x}) = \sum_{i=1}^{s} x'^2_i - \sum_{i=s+1}^{r} x'^2_i .$$

定义5 称可以按公式

$$q(\mathbf{x}) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 \tag{11}$$

计算值的二次型有标准型.

得到的论断表明, 当$\mathfrak{K} = \mathbb{R}$时, 定理4的推论1中的更强的一个不变性成立.

推论1′ 实向量空间V上的所有二次型q均可以归结为标准型.

$\mathbb{R}$上的向量空间V上的二次型q除了秩r之外还有一个数字特性——它的标准型中系数1的个数s. 可以证明, 数s同样不依赖于将q化成标准型的具体的转化方法.

定理5(惯性定律) 设q是$\mathbb{R}$上n维向量空间V上的一个二次型. 那么, 在标准型(11)中的整数r和s, $s \leqslant r \leqslant n$, 都只仅仅依赖于$q$.

证明 r的不变性是已知的, 所以只需证实数s的不变性(与规范基的选择无关). 设另有一个基底, $(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$, 在其上$q$具有标准型

$$q(\mathbf{x}) = (x'_1)^2 + \cdots + (x'_t)^2 - (x'_{t+1})^2 - \cdots - (x'_r)^2, \tag{11'}$$

它有t个正数($\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i = \sum_{i=1}^{n} x'_i\mathbf{e}'_i$), 当$t \neq s$时, 不失一般性, 我们可以认为$t < s$.

研究V的子空间

$$L = \langle \mathbf{e}_1, \cdots, \mathbf{e}_s \rangle_{\mathbb{R}}, \quad L' = \langle \mathbf{e}'_{t+1}, \cdots, \mathbf{e}'_n \rangle_{\mathbb{R}}.$$

因为dim $(L + L')$ ⩽dim $V \leqslant n$, 那么, 据§2定理6, 有

$$\begin{aligned}\dim(L \cap L') &= \dim L + \dim L' - \dim(L + L') \\ &\geqslant s + (n - t) - n = s - t > 0.\end{aligned}$$

可见, 存在一个非零向量$\mathbf{a} \in (L \cap L')$:

$$\mathbf{0} \neq \mathbf{a} = a_1\mathbf{e}_1 + \cdots + a_s\mathbf{e}_s = a'_{t+1}\mathbf{e}'_{t+1} + \cdots + a'_n\mathbf{e}'_n.$$

对照(11)

$$q(\mathbf{a}) = a_1^2 + \cdots + a_s^2 > 0.$$

而同时对照(11′)

$$q(\mathbf{a}) = -(a'_{t+1})^2 - \cdots - (a'_r)^2 \leqslant 0$$

(有可能, $r < n, a'_{t+1} = \cdots = a'_r = 0$). 由得到的这个矛盾可推知只能$s = t$. □

鉴于定理5, 对数字不变性可使用专门的术语.

定义6 *称实二次型的秩是它的***惯性指数**, *数s为***正惯性指数**, $r - s$*是***负惯性指数**. *而***符号差**一词*既可以理解为数对$(s, r - s)$, 也可以理解为正数个数与负数个数之差$2s - r$.*

二次型的惯性定律归功于西尔维斯特(Sylvester, 1814—1897), 起源于力学. 显然, 对于复的二次型$q : V \to \mathbb{C}$, 正负惯性指数失去任何意义, 因为在它的对角形式(10)中做成的所有的λ_i都等于1或者可都等于-1.

8. 正定型与正定矩阵. 仍然设V是个实的向量空间. V上二次型q称为**非退化的**, 如果rank $q = \dim_{\mathbb{R}} V$, 换言之, 它的惯性指数和它的秩一致.

定义7 *非退化的二次型$q : V \to \mathbb{R}$ 称为是***正定的**(**负定的**), *如果$q(\mathbf{x}) > 0(q(\mathbf{x}) < 0)$对任意向量$\mathbf{x} \neq 0$都成立. 二次型$q$称为是***半正定的**(*或***非负定的**), *如果$q(\mathbf{x})\geqslant 0$对所有$\mathbf{x} \in V$都成立. 最后, 二次型q称为是***不定的**, *如果它有时取正值有时取负值.*

特别要注意, 这些概念都与基底的选择没有关系, 相对应的$n = \dim_{\mathbb{R}} V$标准型是:

正定情形, $x_1^2+x_2^2+\cdots+x_n^2$;

负定情形, $-x_1^2-x_2^2-\cdots-x_n^2$;

半正定情形, $x_1^2+x_2^2+\cdots+x_r^2, \quad r\leqslant n$;

不定情形, $r>s>0$, 见(11).

在规范基底$(\mathbf{e}_i)$之下记实的二次型, 显然, $q(\mathbf{x})=\lambda_1x_1^2+\cdots+\lambda_nx_n^2$是正定的, 当且仅当, 所有系数$\lambda_i$都大于零. 要充分注意, $\lambda_i=q(\mathbf{e}_i)$.

与正定的二次型相配极的双线性型也同样称为是**正定的**. 类似的术语可以照搬到矩阵上. 例如, 实对称矩阵F称为是**正定的**, 如果F对应正定的二次型. 但正定矩阵在规范基底之下对应单位矩阵. 对照定理4的推论2, 就有

定理6 任意正定的矩阵F必然形如

$$F={}^tA\cdot A, \tag{12}$$

其中, A是实的非退化矩阵. 反过来也对: 任意一个形如(12)的实矩阵都是正定的.

常常有必要直接利用二次型的矩阵去判断它是不是正定的.

例 设$\varphi(x,y)$是两个实变量的可微函数, 允许在坐标原点的邻域有收敛的泰勒级数分解. φ_x',φ_y'分别表示它对x和y的偏导数. 设点(0, 0)是**临界的**(或者也说是**平稳的**), 即$\varphi_x'(0,0)=0=\varphi_y'(0,0)$, 因此, 在泰勒级数分解中, 从零次项和二次项开始, 即

$$\varphi(x,y)=\varphi(0,\ 0)+\frac{1}{2}\left\{ax^2+2bxy+cy^2\right\}+\cdots$$

这里, $a=\varphi_{xx}''(0,0), b=\varphi_{xy}''(0,\ 0), c=\varphi_{yy}''(0,\ 0)$, 而省略的地方代表更高次的项, 在0的足够小的邻域上, 这些项都可以忽略不计. 因此, 函数φ的值近似等于$\varphi(0,\ 0)$加上$\frac{1}{2}q(\mathbf{v}), \mathbf{v}=x\mathbf{e}_1+y\mathbf{e}_2$, 其中

$$q(\mathbf{v})=ax^2+2bxy+cy^2,$$

这是一个$V=\langle\mathbf{e}_1,\mathbf{e}_2\rangle$上的二次型. 在一般rank q=2的情形, 就称临界点(0, 0)是**非退化的**. 如果q是正定的, 那么, 显然, φ在(0, 0)有相对**极小值**. **极大值**则对应负定型q. 如果二次型q的符号差是(1, 1), 那么, 在(0, 0)既不是极大也不是极小, 此时临界点(0, 0)被称为**鞍点**.

把$q(\mathbf{v})$写成

$$q(\mathbf{v})=a\left(x+\frac{by}{a}\right)^2+\left(c-\frac{b^2}{a}\right)y^2, a\neq 0$$

的形式, 或者, 当$a=0, c\neq 0$时, 用类似的表达, 我们可以看出

$$a>0, \begin{vmatrix} a & b \\ b & c \end{vmatrix}>0$$

是二次型q正定的一个十分简洁的充要条件. 从而也是函数φ在坐标原点邻域极小值存在的充要条件.

在n维情形, 类似于研究上面那些不等式引出来的行列式被称为是矩阵$F(f_{ij})$的**主子式**

$$\Delta_1=f_{11},\Delta_2=\begin{vmatrix} f_{11} & f_{12}\\ f_{21} & f_{22}\end{vmatrix},\cdots,\Delta_k=\begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1k}\\ f_{21} & f_{22} & \cdots & f_{2k}\\ \cdots & \cdots & \cdots & \cdots\\ f_{k1} & f_{k2} & \cdots & f_{kk}\end{vmatrix},\cdots \tag{13}$$

这样一来, $\Delta_n=\det F$. 为了方便, 设$\Delta_0=1$. 主子式的作用可以用一种特殊的把二次型化成规范型的方法加以刻画.

定理7(雅可比方法) *设q是V上以F为矩阵的二次型, F的所有主子式(13)都不等于零. 那么, 必有空间V的基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$, 在此基之下$q(\mathbf{x})$具有规范形式*

$$q(\mathbf{x})=\frac{\Delta_0}{\Delta_1}(x'_1)^2+\frac{\Delta_1}{\Delta_2}(x'_2)^2+\cdots+\frac{\Delta_{n-1}}{\Delta_n}(x'_n)^2. \tag{14}$$

证明 设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是$V$的一个初始的基底, 看$n-1$维子空间$L=\langle\mathbf{e}_1,\cdots,\mathbf{e}_{n-1}\rangle$. 设$\overline{q}=q|_L$ 是q在L上的限制. 型$\overline{q}$的矩阵$\overline{F}$是由F去掉最后一行与最后一列得到的, 故它的主子式是$\overline{\Delta}_1=\Delta_1,\cdots,\overline{\Delta}_{n-1}=\Delta_{n-1}$. 据条件, 它们都不为零. 对$n$用归纳法, 可以断言, 在$L$中可选取基底, 在此基之下. $\overline{q}(\overline{\mathbf{x}})$, $\overline{\mathbf{x}}\in L$具有形式

$$\overline{q}(\overline{\mathbf{x}})=q(\overline{\mathbf{x}})=\frac{\Delta_0}{\Delta_1}(x'_1)^2+\cdots+\frac{\Delta_{n-2}}{\Delta_{n-1}}(x'_{n-1})^2.$$

把这个事实换成与之配极的双线性型f的术语:

$$f(\mathbf{e}'_i,\mathbf{e}'_i)=\frac{\Delta_{i-1}}{\Delta_i},\quad f(\mathbf{e}'_i,\mathbf{e}'_j)=0,\quad 1\leqslant i\neq j\leqslant n-1.$$

n个未知量$x'_1,\cdots,x'_n$的$n-1$个齐次线性方程组

$$f(\mathbf{x},\mathbf{e}'_1)=0,\cdots,f(\mathbf{x},\mathbf{e}'_{n-1})=0,\quad \mathbf{x}\in V,$$

显然在V中必有非零解; $\mathbf{x}=\mathbf{e}'_n$是这个对应解. 容易看出来, 向量组$(\mathbf{e}'_1,\mathbf{e}'_2,\cdots,\mathbf{e}'_n)$构成$V$的一个基底. 因为确定向量$\mathbf{e}'_n$可以精确到纯量因子, 我们用条件规范它, 可以使得由基底$(\mathbf{e}_i)$向基底$(\mathbf{e}'_i)$的转换矩阵A有行列式

$$\det A=(\Delta_n)^{-1}=(\det F)^{-1}.$$

设F'是线性型f在基底$(\mathbf{e}'_i)$之下的矩阵. 那么, 当$i\neq j$时, $f(\mathbf{e}'_i,\mathbf{e}'_j)=0$, 而且

$$\begin{aligned}\frac{f(\mathbf{e}'_n,\mathbf{e}'_n)}{\Delta_{n-1}}&=\frac{\Delta_0}{\Delta_1}\cdot\frac{\Delta_1}{\Delta_2}\cdots\frac{\Delta_{n-2}}{\Delta_{n-1}}\cdot f(\mathbf{e}'_n,\mathbf{e}'_n)=\prod_{i=1}^{n}f(\mathbf{e}'_i,\mathbf{e}'_i)\\&=\det F'=\det({}^tA\cdot F\cdot A)=(\det A)^2\det F=\frac{1}{\Delta_n},\end{aligned}$$

从而

$$f(\mathbf{e}'_n, \mathbf{e}'_n) = \frac{\Delta_{n-1}}{\Delta_n}.$$

二次型q在基底$(\mathbf{e}'_i)$之下表达出来, 就变成了所要求的(14)的形式了. □

容易相信, 矩阵A是三角形的:

$$\begin{aligned}
\mathbf{e}'_1 &= a_{11}\mathbf{e}_1,\\
\mathbf{e}'_2 &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2,\\
&\cdots\cdots\cdots\cdots\\
\mathbf{e}'_n &= a_{n1}\mathbf{e}_1 + a_{n2}\mathbf{e}_2 + \cdots + a_{nn}\mathbf{e}_n,
\end{aligned}$$

但这个事实对我们不是必需的.

推论 矩阵F的所有主子式$\Delta_i, 1 \leqslant i \leqslant n$均不为零, 与$F$对应的二次型$q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$的负惯性指数必与下列的

$$1 = \Delta_0, \Delta_1, \cdots, \Delta_n,$$

变号的个数相同. 如果, 特别地,

$$\Delta_1 > 0, \cdots, \Delta_n > 0,$$

那么, 二次型q就是正定的.

现在, 我们看这个推论的反命题.

定理8(西尔维斯特准则) 实的n维向量空间V上的二次型q是正定的, 当且仅当, 它的矩阵F=(f_{ij})的所有主子式$\Delta_1, \cdots, \Delta_n$是正的.

证明 对照定理7的推论, 不等式$\Delta_i > 0, i = 1, 2, \cdots, n$保障了二次型$q$的正定性. 为了证明反过来的论断, 可像定理7一样对n用归纳法. 看二次型q在$n-1$维子空间$U = \langle \mathbf{e}_1, \cdots, \mathbf{e}_{n-1} \rangle \subset \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle = V$ ($(\mathbf{e}_i)$是V的基底, 此基之下q的矩阵为F)上的限制$\bar{q} = q|_U$.

显然, 二次型$\bar{q}$的矩阵$\overline{F}$的主子式是$\overline{\Delta}_1 = \Delta_1, \cdots, \overline{\Delta}_{n-1} = \Delta_{n-1}$. 因为已假定$q$是正定的, 所以二次型$\bar{q}$也是正定的, 据归纳法假设, $\Delta_1 > 0, \cdots, \Delta_{n-1} > 0$. 剩下的是只要证明$\Delta_n > 0$. 但, 由定理6, 我们知道$F = {}^tAA$, 其中$A$是一个非退化矩阵. 所以,

$$\Delta_n = \det F = \det {}^tA \cdot \det A = (\det A)^2 > 0.$$ □

9. 斜对称二次型的规范型 集中主要精力研究了二次型(同时也就研究了双线性型)之后, 现在遵循定理2, 我们转向空间$\mathcal{L}_2^-(V, \mathfrak{K})$, 也就是斜对称的双线性型. 换言之, 设

$$f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x}), \qquad \forall \mathbf{x}, \mathbf{y} \in V.$$

和对称的双线性型的情形一样, 子空间

$$V_0 = \operatorname{Ker} f = \{\mathbf{v} \in V | f(\mathbf{v}, \mathbf{x}) = 0, \forall \mathbf{x} \in V\}.$$

称为双线性型f的**根**(或**核**). 如果V_1是V_0在V中的任意一个补空间, 那么

$$V = V_0 \oplus V_1,$$

并且限制$f|_{V_1}$必为非退化的斜对称型. 事实上, 如果$\mathbf{a} \in V_1, \mathbf{a} \neq \mathbf{0}$且$f(\mathbf{a}, \mathbf{x}_1) = 0$对所有的$\mathbf{x}_1 \in V_1$都成立, 那么, 对任意向量$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1 \in V \quad (\mathbf{x}_0 \in V_0)$, 有

$$f(\mathbf{a}, \mathbf{x}) = f(\mathbf{a}, \mathbf{x}_0 + \mathbf{x}_1) = f(\mathbf{a}, \mathbf{x}_0) + f(\mathbf{a}, \mathbf{x}_1) = -f(\mathbf{x}_0, \mathbf{a}) = 0$$

(这里, 我们利用了型f的斜对称性), 这与V_0的定义矛盾.

从而可将对f的研究归结为非退化情形. 一开始就可以认为$f : V \times V \to \mathcal{R}$是个非退化的双线性型. 设

$$V = \langle \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n \rangle, \qquad \mathbf{x} = \sum_i x_i \mathbf{e}_i, \mathbf{y} = \sum_j y_j \mathbf{e}_j,$$

那么

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{n} f_{ij} x_i y_j = {}^t X F Y, \quad f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j),$$

其中$F = (f_{ij})$是斜对称矩阵: ${}^tF + F = 0$, 也就是

$$f(\mathbf{x}, \mathbf{y}) = \sum_{1 \leqslant i < j \leqslant n} f_{ij}(x_i y_j - x_j y_i). \tag{15}$$

由[BA I]第3章§2可知, 对于n阶斜对称矩阵的行列式, 关系式$\{1 + (-1)^{n-1}\} \cdot \det F = 0$成立. 所以, 要不等式$\det F \neq 0$ (二次型f非退化的条件)成立只有n为偶数时才有可能. 我们可以用另外的方法得到这个结论, 顺便把f化成规范型.

为了这个目的, 引进V中**双曲面**(或**辛平面**)W的概念, 把W理解为有条件$f|_W \neq 0$的任意的二维子空间. 这种子空间可以这样描述: 对任意向量$\mathbf{e'}_1 \neq \mathbf{0}$, 必存在一向量$\mathbf{e}'_2$使得$f(\mathbf{e}'_1, \mathbf{e}'_2) \neq 0$. 用所得到的纯量去除$\mathbf{e}'_2$, 我们就可以认为$f(\mathbf{e}'_1, \mathbf{e}'_2) = 1$, 当然还有$f(\mathbf{e}'_1, \mathbf{e}'_1) = 0 = f(\mathbf{e}'_2, \mathbf{e}'_2)$.

定理9 *设V是个实空间, 其上有一个非退化的斜对称的双线性型f. 那么, $\dim V = 2m$且V是m个对于f而言两两相互斜正交的双曲面的直和.*

证明 对n=$\dim V$用归纳法. 由上面所做的说明, 可得到一个双曲面$W = \langle \mathbf{e}'_1, \mathbf{e}'_2 \rangle \subset V$. 如果$n > 2$, 看斜正交的补

$$W^{\perp} = \langle \mathbf{x} \in V | f(\mathbf{e}'_i, \mathbf{x}) = 0, i = 1, 2 \rangle.$$

把$\mathbf{e}'_1, \mathbf{e}'_2$扩充成空间$V$的(打上一撇的)基底. 设

$$V = \langle \mathbf{e}'_1, \mathbf{e}'_2, \cdots, \mathbf{e}'_n \rangle, \quad \mathbf{x} = x'_1 \mathbf{e}'_1 + \cdots + x'_n \mathbf{e}'_n.$$

那么,

$$f(\mathbf{e}_1', \mathbf{x}) = f_{12}'x_2' + f_{13}'x_3 + \cdots + f_{1n}'x_n' = 0,$$
$$f(\mathbf{e}_2', \mathbf{x}) = f_{21}'x_1' + f_{23}'x_3' + \cdots + f_{2n}'x_n' = 0$$

是秩为2的线性方程组, 因为矩阵F的行是线性无关的. 可以知道, 这个线性方程组的解空间$\langle \mathbf{e}_1', \mathbf{e}_2' \rangle^\perp$的维数是$n-2$. 因为

$$\langle \mathbf{e}_1', \mathbf{e}_2' \rangle \cap \langle \mathbf{e}_1', \mathbf{e}_2' \rangle^\perp \subseteq \mathrm{Ker} f = 0,$$

所以可得到分解

$$V = \langle \mathbf{e}_1', \mathbf{e}_2' \rangle \oplus \langle \mathbf{e}_1', \mathbf{e}_2' \rangle^\perp,$$

而且f在$\langle \mathbf{e}_1', \mathbf{e}_2' \rangle^\perp$上的限制是个非退化的斜对称型. 按归纳法假设, $\langle \mathbf{e}_1', \mathbf{e}_2' \rangle^\perp$是个偶数维空间而且是两两相互斜正交的双曲面的直和. 可知, 有某个整数m使$n = \dim V = 2m$而且V有一个基底$(\mathbf{e}_i'')$且$\mathbf{e}_1'' = \mathbf{e}_1', \mathbf{e}_2'' = \mathbf{e}_2'$, 使得

$$V = \langle \mathbf{e}_1'', \mathbf{e}_2'' \rangle \oplus \langle \mathbf{e}_3'', \mathbf{e}_4'' \rangle \oplus \cdots \oplus \langle \mathbf{e}_{2m-1}'', \mathbf{e}_{2m}'' \rangle,$$
$$f(\alpha \mathbf{e}_{2i-1}'' + \beta \mathbf{e}_{2i}'', \gamma \mathbf{e}_{2j-1}'' + \delta \mathbf{e}_{2j}'') = 0, \quad i \neq j,$$
$$f(\mathbf{e}_{2i-1}'', \mathbf{e}_{2i}'') = 1. \qquad \square$$

推论 任意非退化的斜对称的$2m \times 2m$阶的矩阵必然合同于

$$J = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

也就是, 可以找到非退化的矩阵A使${}^tAFA = J$.

如果利用双线性型的矩阵在向新的基底转换中的变化定律, 证明可以从定理9直接推出来.

说明 定理及其推论在任意特征$\neq 2$的域上都成立. 常常选择

$$J_0 = \begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix},$$

作为标准的斜对称的矩阵. 只要对换基底中向量的顺序就可以把J化成J_0.

如果要讨论具体地将斜对称型(15)化成规范型的问题, 还要再用拉格朗日方法. 先设$f \neq 0$(相反情形, 就没有什么可做了), 并且必要时交换基底中向量顺序, 从而转向$f_{12} \neq 0$的情形. 在(15)中提出所有带有变量x_1或者y_1的被加项:

$$x_1(f_{12}y_2 + \cdots + f_{1n}y_n) - (f_{12}x_2 + \cdots + f_{1n}x_n)y_1.$$

从新的变量

$$x_2' = f_{12}x_2 + \cdots + f_{1n}x_n, \quad y_2' = f_{12}y_2 + \cdots + f_{1n}y_n$$

$(x_1, y_1, x_3, y_3, \cdots, x_n, y_n$仍然不变)开始计算含有$x_2', y_2'$的被加项的过程

$$\begin{aligned} & f(\mathbf{x}, \mathbf{y}) \\ = & (x_1 + f_{32}'x_3 + \cdots + f_{n2}'x_n)y_2' - x_2'(y_1 + f_{32}'y_3 + \cdots + f_{n2}'y_n) + \cdots \\ = & x_1'y_2' - x_2'y_1' + \cdots, \end{aligned}$$

其中

$$x_1' = x_1 + f_{32}'x_3 + \cdots + f_{n2}'x_n, \quad y_1' = y_1 + f_{32}'y_3 + \cdots + f_{n2}'y_n,$$

上面省略的被加项只包含$x_3, y_3, \cdots, x_n, y_n$. 如果它们不等于零, 那么就用同样的方法处置它们. 最后, 型f的项被化成了规范型

$$f(\mathbf{x}, \mathbf{y}) = (x_1'y_2' - x_2'y_1') + \cdots + (x_{2m-1}'y_{2m}' - x_{2m}'y_{2m-1}'). \tag{16}$$

10. 普法夫型 据定理9的推论, 对每个非退化的斜对称矩阵F 都必有矩阵A 使之

$${}^tAFA = \begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix} = J_0.$$

从而$(\det A)^2 \det F = 1$, 即$\det F$是基础域$\mathfrak{K}$上的一个平方数. 这个情况引出一个想法要研究: 以斜对称矩阵$T = (t_{ij}), -t_{ij} = t_{ji}, i < j$的$n(n-1)/2$个无关变量构成的多项式环

$$\mathbb{Z}[t] = \mathbb{Z}[t_{12}, t_{13}, \cdots, t_{n-1,n}]$$

的商域

$$\mathfrak{K} = \mathbb{Q}(t) = \mathbb{Q}(t_{12}, t_{13}, \cdots, t_{n-1,n})$$

我们知道, $\det T$是域$\mathbb{Q}(t)$上的一个二次型. 而另一方面, $\det T$也是$\mathbb{Z}[t]$中的一个多项式. 要明白(这里我们模模糊糊地应用$\mathbb{Z}[t]$中乘积分解的唯一性; 见[BA Ⅰ]和[BA Ⅲ]), $\det T$是$\mathbb{Z}[t]$中某个多项式的平方:

$$\det T = P_n(t)^2.$$

将$P_n(t) = P_n(t_{12}, t_{13}, \cdots)$规范化, 使其对某个控制值$t_{ij}^0 = 0, \pm 1$有$P_n(t_{12}^0, t_{13}^0, \cdots) = 1$, 而$T_0 = (t_{ij}) = J_0$. 在这种规范之下, 得到唯一确定的多项式$\mathrm{Pf}_n(t)$, 一般地, 称它为$n$维的**普法夫**. 例如,

$$\mathrm{Pf}_2(t) = t, \quad \mathrm{Pf}_4(t) = t_{12}t_{34} - t_{13}t_{24} + t_{14}t_{23},$$

它容易计算矩阵

$$\begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -t_{12} & -t_{13} & -t_{14} \\ t_{12} & 0 & -t_{23} & -t_{24} \\ t_{13} & t_{23} & 0 & -t_{34} \\ t_{14} & t_{24} & t_{34} & 0 \end{pmatrix}$$

的行列式.

将$\mathrm{Pf}(F)$理解为用斜对称矩阵F的系数f_{ij}代替到t_{ij}位置上得到的$\mathrm{Pf}_n(f_{ij})$ (把各处的$\mathbb{Q}$换成素域Z_p, 可将我们的论断推广到任意特征数的域上), 有

定理10 如果F是任意一个$n \times n$阶斜对称矩阵, 那么,

$$\det F = \mathrm{Pf}(F)^2.$$

进一步, 还有

$$\mathrm{Pf}({}^tAFA) = \det A \cdot \mathrm{Pf}(F)$$

证明 关系式$\det F = \mathrm{Pf}(F)^2$表达出我们已经知道的斜对称矩阵和普法夫多项式的性质. 进而, 设$U = (u_{ij})$是任意$n \times n$个代数无关的u_{ij}的矩阵, T是前面看到过的斜对称矩阵. 那么,

$$\mathrm{Pf}({}^tUTU)^2 = \det({}^tUTU) = (\det U)^2 \det T = (\det U)^2 \mathrm{Pf}(T)^2,$$

从而

$$\mathrm{Pf}({}^tUTU) = \pm(\det U)\mathrm{Pf}(T).$$

如果用那样的值来代替u_{ij}和t_{ij}使得U是个归一矩阵而T是个标准的斜对称矩阵, 那么, 左边$\mathrm{Pf}(J_0) = 1$, 而右边是$\pm 1 \cdot \mathrm{Pf}(J_0)$, 也就是应该取+号. 这说明, 对特殊的矩阵$U = A, T = F$, 要证明的等式是成立的. □

剩下来需要说明的是, 对于$T = (t_{ij})_1^{2m}, {}^tT = T$, 普法夫式$\mathrm{Pf}(T)$是个泛多项式, 它是齐$m$次型, 它的系数是整数或者属于素域.

习　　题

1. 设$\Delta_1, \cdots, \Delta_n = F$是矩阵$F$对应的实二次型$q$的主子式. 证明, 只有当$(-1)^k\Delta_k > 0$, 对所有$k = 1, 2, \cdots, n$都成立, q和F才能是负定的.

2. 举出反例:

(1) 正定矩阵$A = (a_{ij})$, 但有某个对(i, j)使得$a_{ij} < 0$;

(2) 矩阵$A = (a_{ij})$, 对所有的指标都有$a_{ij} > 0$, 但A却不是正定的.

3. 找出所有$\lambda, \mu \in \mathbb{R}$, 使得矩阵

$$\begin{pmatrix} 1 & \lambda & \lambda \\ \lambda & 1 & \lambda \\ \lambda & \lambda & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & \mu \\ 1 & \mu & 1 \\ \mu & 1 & 1 \end{pmatrix}$$

是正定的.

4. 设$\mathbf{x} = [x_1, x_2, x_3] \in \mathbb{C}^3, Q(\mathbf{x}) = x_1^3 + x_2^3 + x_3^3 - 3x_1x_2x_3$, ε是三次本原根, 利用表示式

$$Q(\mathbf{x}) = (x_1 + x_2 + x_3)(x_1 + \varepsilon x_2 + \varepsilon^2 x_3)(x_1 + \varepsilon^2 x_2 + \varepsilon x_3),$$

证明$Q(\mathbf{x})Q(\mathbf{y}) = Q(\mathbf{z})$, 其中$\mathbf{z} = [z_1, z_2, z_3]$, $z_i = z_i(\mathbf{x}, \mathbf{y}) = \sum\limits_{j,k} a_{jk}^{(i)} x_j y_k$是对称的双线性型, 找出它们的显式.

5. 设A是任意一个实对称矩阵, $\varepsilon = \varepsilon(A)$是个充分小的实数. 证明, 矩阵$B = E + \varepsilon A$是正定的.

第2章 线性算子

照例, 不仅要研究向量空间本身, 同时还要研究向量空间的线性映射. 在$\mathbb{R}^3_\Phi$中的旋转、反射、同位相似以及分析学中的微分算子和积分算子都可以充当这方面的例子. 开始, 我们把精力集中在线性映射的最一般的性质上面.

§1 向量空间的线性映射

1. 线性映射语言 如我们所知(见[BA Ⅰ] 第2章§3), 每个$m\times n$的矩阵A都对应一个线性映射φ_A: $\mathbb{R}^n\to\mathbb{R}^m$. 仿照它的性质, 引入下面一般的

定义1 设V, W是同一个域$\mathfrak{K}$上的n维、m维向量空间. 映射$f:V\to W$称为是**线性**的, 如果

$$f(\mathbf{x}+\mathbf{y})=f(\mathbf{x})+f(\mathbf{y}),\quad f(\lambda\mathbf{x})=\lambda f(\mathbf{x}).$$

换句话说, $f(\alpha\mathbf{x}+\beta\mathbf{y})=\alpha f(\mathbf{x})+\beta f(\mathbf{y})$.在第1章已经研究过的线性函数$f:V\to\mathfrak{K}$的概念是线性映射的一个特殊的类型.

所有$V\to W$的线性映射的集合, 用符号$\mathcal{L}(V,\,W)$(或者Hom $(V,\,W)$)表示, 它是个具有自然的映射加法和纯量乘法的向量空间: 如果$f,g\in\mathcal{L}(V,W)$且$\nu,\mu\in\mathfrak{K}$, 那么, 按定义

$$(\nu f+\mu g)(\mathbf{x})=\nu f(\mathbf{x})+\mu g(\mathbf{x}).$$

可以直接验证, 向量空间的所有公理(见第1章, §1)针对$\mathcal{L}(V,\,W)$都被满足.

和任意一个线性映射$f:V\to W$联系在一起有两个子空间, 它的核

$$\mathrm{Ker}\ f=\{\mathbf{v}\in V|f(\mathbf{v})=0\}$$

和像

$$\mathrm{Im}f=\left\{\mathbf{w}\in W|\mathbf{w}=f(\mathbf{v}),\text{对某个}\mathbf{v}\in V\right\}.$$

核和像对我们来说并不是一个新的概念. 但是, 现在要特别强调, 它们分别是V和W的子空间(对于Ker f容易检验, 而对Imf的检验将在下面进行). 对任意子空间$U \subset V$, 我们约定, 记$f(U) = \{f(\mathbf{u})|\mathbf{u} \in U\}$.如果, $\mathbf{u}_1, \mathbf{u}_2 \in U, \nu_1, \nu_2 \in \mathfrak{K}$, 那么

$$\nu_1 f(\mathbf{u}_1) + \nu_2 f(\mathbf{u}_2) = f(\nu_1 \mathbf{u}_1 + \nu_2 \mathbf{u}_2) \in f(U),$$

因为$\nu_1 \mathbf{u}_1 + \nu_2 \mathbf{u}_2 \in U$.所以, $f(U)$是W的一个子空间. 特别地, 这导致$f(V)$=Im f.

注意, f的单射性等价于等式Ker $f = \{\mathbf{0}\}$. 事实上, 在$\mathbf{x} \neq \mathbf{y}, f(\mathbf{x}) = f(\mathbf{y})$的情形, 就有$\mathbf{0} \neq \mathbf{x} - \mathbf{y} \in \text{Ker} f$. 反之, 如果$\mathbf{0} \neq \mathbf{x} \in \text{Ker} f$, 那么$f(\mathbf{x}) = 0 = f(\mathbf{0})$.

定理1 设$f : V \to W$是个线性映射. 如果$U = \langle \mathbf{e}_1, \cdots, \mathbf{e}_s \rangle \subset V$, 那么$f(U) = \langle f(\mathbf{e}_1), \cdots, f(\mathbf{e}_s) \rangle$. 特别地, 有

$$\dim f(U) \leqslant \dim U.$$

证明 据条件, 任意$\mathbf{u} \in U$可记成$\mathbf{u} = \alpha_1 \mathbf{e}_1 + \cdots + \alpha_s \mathbf{e}_s$.因此, $f(\mathbf{u}) = \alpha_1 f(\mathbf{e}_1) + \cdots + \alpha_s f(\mathbf{e}_s)$, 这就意味着, $f(U) = \langle f(\mathbf{e}_1), \cdots, f(\mathbf{e}_s) \rangle$. 当$(\mathbf{e}_1, \cdots, \mathbf{e}_s)$是$U$的基底的时候, 一般说来, $(f(\mathbf{e}_1), \cdots, f(\mathbf{e}_s))$未必是$f(U)$的基底. 同时, $\dim f(U) \leqslant s = \dim U$.完全可能发生$U \subseteq \text{Ker } f$与$f(U) = \{\mathbf{0}\}$这样的事情. □

2. 用矩阵给定线性映射 设分别给定了向量空间V, W的基底$(\mathbf{v}_1, \cdots, \mathbf{v}_n)$和 $(\mathbf{w}_1, \cdots, \mathbf{w}_m)$. 像Im $f \subset W$的每个向量都是向量组

$$\begin{aligned} f(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ f(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m. \end{aligned} \tag{1}$$

的一个线性组合. 反之, 给空间W的一组向量派定任务$\mathbf{w}_1' = f(\mathbf{v}_1), \cdots, \mathbf{w}_n{}' = f(\mathbf{v}_n)$就完全确定了一个线性映射$f$; 任意向量$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$, 完全由$f(\mathbf{v}) = \alpha_1 \mathbf{w}_1' + \cdots + \alpha_n \mathbf{w}_n'$给出. 如果$\mathbf{v}' = \alpha_1' \mathbf{v}_1 + \cdots + \alpha_n' \mathbf{v}_n$, 那么

$$\begin{aligned} f(\nu\mathbf{v} + \nu'\mathbf{v}') &= f((\nu\alpha_1 + \nu'\alpha_1')\mathbf{v}_1 + \cdots + (\nu\alpha_n + \nu'\alpha_n')\mathbf{v}_n) \\ &= (\nu\alpha_1 + \alpha_1')\mathbf{w}_1' + \cdots + (\nu\alpha_n + \nu'\alpha_n')\mathbf{w}_n' \\ &= \nu(\alpha_1\mathbf{w}_1' + \cdots + \alpha_n\mathbf{w}_n') + \nu'(\alpha_1'\mathbf{w}_1' + \cdots + \alpha_n'\mathbf{w}_n') = \nu f(\mathbf{v}) + \nu' f(\mathbf{v}'). \end{aligned}$$

矩阵

$$M_f = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{2}$$

称为**线性映射**$f : V \to W$**相对于空间**V, W**的基底**$(\mathbf{v}_1, \cdots, \mathbf{v}_n)$, $(\mathbf{w}_1, \cdots, \mathbf{w}_n)$(或基底$(\mathbf{v}_j), (\mathbf{w}_i)$)的**矩阵**. 不同的线性映射对应不同的矩阵.

注意, 向量$f(\mathbf{v}_j)$的坐标组成矩阵M_f的第j列. 因此, $\mathrm{rank}\{f(\mathbf{v}_1),\cdots,f(\mathbf{v}_n)\}=\mathrm{rank}\,M_f$. 因为, 总有$\mathrm{rank}\{f(\mathbf{v}_1),\cdots,f(\mathbf{v}_n)\}=\dim<f(\mathbf{v}_1),\cdots,f(\mathbf{v}_n)\}_{\mathfrak{K}}=\dim\mathrm{Im}\,f$, 所以, $\dim\mathrm{Im}\,f=\mathrm{rank}\,M_f$.

定义2 把子空间$\mathrm{Im}\,f$的维数同时称为线性映射f的**秩**($\mathrm{rank}\,f$).

定理2 i) 设$V=\langle\mathbf{v}_1,\cdots,\mathbf{v}_n\rangle, W=\langle\mathbf{w}_1,\cdots,\mathbf{w}_m\rangle$是带有固定基底的两个向量空间. 那么, 在$V$到$W$的线性映射和系数在基础域$\mathfrak{K}$上的$m\times n$阶矩阵之间存在一个一一对应.

ii) 任意向量组$\mathbf{w}'_1,\cdots,\mathbf{w}'_n\in W$都对应唯一一个线性映射$f:V\to W$, 满足$f(\mathbf{v}_i)=\mathbf{w}'_i, 1\leqslant i\leqslant n$.

iii) 线性映射$f:V\to W$的秩与它对应的矩阵M_f (在V和W的任意基底之下)的秩一致.

正如我们已经知道的那样, 系数在基础域$\mathfrak{K}$上的所有$m\times n$阶矩阵构成$\mathfrak{K}$上mn维向量空间, 它以第i行第j列元素为1, 其余各处均为0的矩阵E_{ij}为基底向量. 因而等式

$$\dim\mathcal{L}(V,W)=(\dim V)(\dim W).$$

成立.

矩阵和线性映射之间的对应可用于已知的关于矩阵乘积的秩的命题(见[BA Ⅰ]第2章§3)的新的证明. 用映射的语言, 它对应

定理3 设$f\circ g$是个映射的复合

$$U\xrightarrow{g}V\xrightarrow{f}W.$$

那么,

i) $\dim\mathrm{Im}(f\circ g)\leqslant\dim\mathrm{Im}\,f$.

ii) $\dim\mathrm{Im}(f\circ g)\leqslant\dim\mathrm{Im}\,g$.

证明 不等式i)是显然的, 因为$\mathrm{Im}(f\circ g)\subset\mathrm{Im}\,f$. 为了证明ii), 我们要注意$\mathrm{Im}(f\circ g)=f(\mathrm{Im}\,g)$. 因为对任意线性映射$h:U\to W$都有$\dim(\mathrm{Im}\,h)\leqslant\dim U$, 故不等式ii)显然成立. □

设$\mathbf{x}=\sum\limits_{j=1}^{n}x_j\mathbf{v}_j$是$V$的向量, $\mathbf{y}=f(\mathbf{x})=\sum\limits_{i=1}^{m}y_i\mathbf{w}_i$是它在线性映射$f:V\to W$之下的像. f在已指定的基底$(\mathbf{v}_j)$, $(\mathbf{w}_i)$之下对应于形式(2)的矩阵为M_f. 那么, 按规则(1)对应, 有

$$f(\mathbf{x})=\sum_{j=1}^{n}x_jf(\mathbf{v}_j)=\sum_{j=1}^{n}x_j\left(\sum_{i=1}^{m}a_{ij}\mathbf{w}_i\right)=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}x_j\right)\mathbf{w}_i=\sum_{i=1}^{m}y_i\mathbf{w}_i.$$

从而$y_i=\sum\limits_{j=1}^{n}a_{ij}x_j,\quad 1\leqslant i\leqslant m$, 或者, 简单地

$$Y=M_f\cdot X,\tag{3}$$

其中$X=[x_1,\cdots,x_n]$, $Y=[y_1,\cdots,y_m]$是向量$\mathbf{x}\in V$和$\mathbf{y}=f(\mathbf{x})\in W$的向量坐标列. 在(3)上我们看到了映射$f$对应的坐标间的线性映射(按思维惯例).

设f, g都是V到W的线性映射. 在固定的基底$(\mathbf{v}_j)$, $(\mathbf{w}_i)$之下, 实际上就导出由$\mathfrak{K}^n$到$\mathfrak{K}^m$的线性映射$f: X\to M_fX, g: X\to M_gX$. 与我们已经给出过的这些表示([BA Ⅰ]第2章§3)相对照, 可以看出, 线性映射$\nu f+\mu g$ (见第1目)对应矩阵

$$M_{\nu f+\mu g}=\nu M_f+\mu M_g.$$

类似地, 线性映射$U\xrightarrow{g}V\xrightarrow{f}W$的复合$f\circ g$在$U$, V, W的固定的基底之下对应矩阵

$$M_{f\circ g}=M_fM_g.$$

再次强调了线性映射与矩阵之间对应的完全一致性, 我们只是回顾了已知的事实.

3. 核与像的维数 下面的定理真确:

定理4 *设V是域$\mathfrak{K}$上的有限维向量空间, $f: V\to W$是个线性映射, 那么*Ker$f$ *和*Im f*都是有限维的, 且*

$$\dim\ \mathrm{Ker}\ f+\dim\ \mathrm{Im}\ f=\dim V.$$

证明 (请与[BA Ⅰ]第2章§3的类似证明相比较). 因为Ker$f\subset V$, 所以dim Ker$f\leqslant\dim V<\infty$. 在核Ker f中选择一个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_k)$并将其扩展成与第1章§2定理3相应的空间$V$的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_k,\mathbf{e}_{k+1},\cdots,\mathbf{e}_n)$. Im f中的任意一个向量必形如

$$f\left(\sum_{i=1}^{n}\alpha_i\mathbf{e}_i\right)=\sum_{i=k+1}^{n}\alpha_i f(\mathbf{e}_i),\quad \alpha_i\in\mathfrak{K},$$

也就是说, 向量$f(\mathbf{e}_{k+1}),\cdots,f(\mathbf{e}_n)$生成Im f. 剩下来只需指明, 这些向量是线性无关的.

设$\sum\limits_{i=k+1}^{n}\lambda_i f(\mathbf{e}_i)=0$. 那么$f(\sum\limits_{i=k+1}^{n}\lambda_i\mathbf{e}_i)=0$. 这意味着, $\sum\limits_{i=k+1}^{n}\lambda_i\mathbf{e}_i\in\mathrm{Ker}\ f$, 即$\sum\limits_{i=k+1}^{n}\lambda_i\mathbf{e}_i=\sum\limits_{j=1}^{k}\lambda_j\mathbf{e}_j$. 但是, 基底向量$\mathbf{e}_1,\cdots,\mathbf{e}_n$ 之间的线性关系式应该是平凡的. 推出结论, $\lambda_{k+1}=\cdots=\lambda_n=0$. 从而, 向量$f(\mathbf{e}_{k+1}),\cdots,f(\mathbf{e}_n)$线性无关且dim Im $f=n-k$. □

推论 *在*dim $f<\infty$*情形, 线性映射f: $V\to W$的下列性质是等价的:*

i) *f是单射;*

ii) dim V=dim Im f.

证明 据定理, dim V = dim Im f当且仅当dim Ker $f=0$, 也就是Ker $f=\{\mathbf{0}\}$. 而我们已经看到, 只有f是单射的情形, f的核为零. □

说明 如果$\dim V = \dim W$ 且$f : V \to W$是个线性映射, 那么, 如同已确立的推论一样, 由内射性(此时$\mathrm{Ker}\ f = \{\mathbf{0}\}$)或者满射性($\mathrm{Im}\ f = W$)都可以推出双射性, 这时$f$是同构映射.

习 题

1. 把坐标列$\mathbf{x} = [x_1, x_2, x_3, x_4]$记成矩阵$\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \in M_2(\mathfrak{K})$形式, 再取一个固定的矩阵$A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \in M_2(\mathfrak{K})$, 定义两个线性映射

$$f_L : X \mapsto A\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = \begin{pmatrix} x_1' & x_2' \\ x_3' & x_4' \end{pmatrix} = X',$$

$$f_R : X \mapsto \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} A = \begin{pmatrix} x_1'' & x_2'' \\ x_3'' & x_4'' \end{pmatrix} = X'',$$

它们分别对应矩阵M_{f_L}和M_{f_R}.

试验证

$$M_{f_L} = \begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \\ a_3 & 0 & a_4 & 0 \\ 0 & a_3 & 0 & a_4 \end{pmatrix}, M_{f_R} = \begin{pmatrix} a_1 & a_3 & 0 & 0 \\ a_2 & a_4 & 0 & 0 \\ 0 & 0 & a_1 & a_3 \\ 0 & 0 & a_2 & a_4 \end{pmatrix}.$$

2. 验证下面映射的线性性:

(1) V是个向量空间, $W = V/L$是V的商空间, f把每个向量$\mathbf{x} \in V$映成陪集$\bar{\mathbf{x}}=\mathbf{x}+L$;

(2) $f : P_n \to P_n$是个映射, 其定义按规则$f(u(t)) = tu'(t) - u(t)$.

找出$\mathrm{Ker}\ f$并计算出$\mathrm{rank}\ f$;

(3) 指明, 由非退化矩阵$C \in M_n(\mathfrak{K})$确定的映射$f_C : X \mapsto C^{-1}XC$在$M_n(\mathfrak{K})$上是线性的而且具有性质$f_C(XY) = f_C(X)f_C(Y)$.

§2 线性算子代数

1. 定义与例子 暂且假定基础域$\mathfrak{K}$是任意的. 在$V = W$的情形, 向量空间$\mathcal{L}(V, W)$ 现在可以自然地记成$\mathcal{L}(V)$(同样也用符号$\mathrm{End}(V)$). 它的向量通常称为**线性算子**或**线性变换**. 鉴于术语"线性变换"的多义性(宁愿把它和向量坐标联系起来, 而不把它与向量空间联系在一起), 我们更喜欢它们中的第一个. 今后, 线性算子将用拉丁字母$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \cdots$表示, 而在向量空间$V$的基底$(\mathbf{e}_i)$之下对应的矩阵用字母$A$, B, C, $D, \cdots$ 表示. 而在另外的基底(打了撇的)$(\mathbf{e}_i')$之下, 算子$\mathcal{A}$, $\mathcal{B}$, $\cdots$ 将对应A', B', $\cdots$. 总是用$\mathcal{E}=\mathrm{Id}$ 和$E=(\delta_{ij})$表示恒等(单位)映射$\mathbf{x}\mapsto\mathbf{x}$. 按惯例, 算子$\mathcal{A}$作用在 $\mathbf{x}$上的结果简单地写成$\mathcal{A}\mathbf{x}$(代替$\mathcal{A}(\mathbf{x})$).

线性算子$\mathcal{B}$ 称为是$\mathcal{A}$ 的**逆算子**, 如果$\mathcal{AB}=\mathcal{BA}=\mathcal{E}$. 与已知的一般结果相对应(见[BA I]第1章§5), 算子$\mathcal{A}$的逆算子如果存在的话, 必唯一确定; 可用符

号$\mathcal{A}^{-1}$记之. 据§1定理4, $\mathcal{A}^{-1}$存在等价于Ker $\mathcal{A}$=0或者dim V=dim Im $\mathcal{A}$. 在一般情形, dim Ker $\mathcal{A}$称为算子$\mathcal{A}$的亏数. 可见, 算子$\mathcal{A}$的亏数为零, 它们就是可逆算子. 例如, 如果方程$\mathcal{A}\mathbf{x}=\mathbf{b}$对任意$\mathbf{b}$都有解, 那么$\mathcal{A}$就有逆算子. 再次记起, rank $\mathcal{A}$=dim Im $\mathcal{A}$=dim V−dim Ker $\mathcal{A}$就是算子$\mathcal{A}$的秩, 用矩阵语言定义的所有的这些概念和条件我们早就熟悉了. 但要特别地承认, 线性的概念更为基本: 它与选择何种基底没有关系.

引入一些线性算子的例子.

例1 零算子$\mathcal{O}$把每个向量$\mathbf{v} \in V$都变成零: rank $\mathcal{O}$=0.

例2 相似算子$\mathcal{A}: \mathcal{A}\mathbf{x} = \lambda\mathbf{x}$ (λ是一个固定的纯量).

例3 平面$\mathbb{R}^2$简单地整体旋转α角的旋转算子$\mathcal{A}$, 把$\mathbb{R}^2$看成是以$\{1, i\}$为基底的平面, 将其整体旋转α角. 于是, 显然$\mathcal{A}: z = x + iy \mapsto e^{i\alpha}z$是这样一个算子即用数$e^{i\alpha}$乘, 且$\mathcal{A} \cdot i = -\sin\alpha + i\cos\alpha$. 所以, 在基底$(1, i)$之下算子$\mathcal{A}$的矩阵是

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

例4 $V = U \oplus W$是子空间的直和. 如果$\mathbf{x} = \mathbf{x}_U + \mathbf{x}_W$是以$\mathbf{x}_U$和$\mathbf{x}_W$为分量的向量分解, 且$\mathcal{P}\mathbf{x} = \mathbf{x}_U$, 那么, 称$\mathcal{P}$为**投影算子**或在子空间$U$平行于$W$的**投影**(或沿着$W$的投影), 注意$\mathcal{P}^2 = \mathcal{P}$.

例5 设$P_n = \langle 1, t, \cdots, t^{n-1}\rangle$是域$\mathfrak{K}$上次数$\leqslant n-1$的多项式的空间, $\mathcal{D}_t = d/dt$是按t求导的微分算子: $\mathcal{D}_t \cdot f(t) = f'(t)$.

应该防止在解释公式

$$\dim \text{Ker } \mathcal{A} + \dim \text{Im } \mathcal{A} = \dim V \tag{1}$$

时可能出现的错误(见§1定理4的公式), 从这里不能总是得出$V = \text{Ker } \mathcal{A} + \text{Im } \mathcal{A}$. 如例5中的算子$\mathcal{D}$表明.

$$\text{Ker } \mathcal{D} = \langle 1\rangle = \mathfrak{K} \cdot 1 \subset \langle 1, t, \cdots, t^{n-2}\rangle = \text{Im } \mathcal{D}.$$

2. 算子代数 我们已经知道, 向量空间V上的所有线性算子$\mathcal{A}$的集合$\mathcal{L}(V)$本身是个维数为

$$\dim \mathcal{L}(V) = (\dim V)^2 \tag{2}$$

的向量空间. 一个线性算子由自己在每个元素$\mathbf{x} \in V$上的作用完全确定. 回顾在[BA I]第1章、第2章叙述的映射复合的原则, 我们令

$$(\mathcal{A} + \mathcal{B})\mathbf{x} = \mathcal{A}\mathbf{x} + \mathcal{B}\mathbf{x}, \quad (\lambda\mathcal{A})\mathbf{x} = \lambda(\mathcal{A}\mathbf{x}), \quad (\mathcal{A}\mathcal{B})\mathbf{x} = \mathcal{A}(\mathcal{B}\mathbf{x})$$

(这样一来, 复合$\mathcal{A} \circ \mathcal{B}$可被简单地表达成$\mathcal{A}$的后面直接写$\mathcal{B}$), 由这个定理可以直接计

算出

$$\begin{aligned}
&\alpha(\mathcal{A}+\mathcal{B})=\alpha\mathcal{A}+\alpha\mathcal{B},\\
&(\alpha+\beta)\mathcal{A}=\alpha\mathcal{A}+\beta\mathcal{A},\\
&(\alpha\beta)\mathcal{A}=\alpha(\beta\mathcal{A}),\\
&1\cdot\mathcal{A}=\mathcal{A},
\end{aligned}\tag{3'}$$

$$\begin{aligned}
&\mathcal{A}(\mathcal{B}\mathcal{C})=(\mathcal{A}\mathcal{B})\mathcal{C}(\text{结合律}),\\
&\mathcal{A}(\mathcal{B}+\mathcal{C})=\mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C},\\
&(\mathcal{A}+\mathcal{B})\mathcal{C}=\mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C}(\text{分配律});
\end{aligned}\tag{3''}$$

$$\lambda(\mathcal{A}\mathcal{B})=(\lambda\mathcal{A})\mathcal{B}=\mathcal{A}(\lambda\mathcal{B}).\tag{3'''}$$

我们看到, 线性映射的集合$\mathcal{L}(V)$同时也是个域$\mathfrak{K}$上的向量空间(最前边的(3′)的四个关系式), 也同时还是个结合环(由(3″)的三个关系式推出); 最后的混合型关系式(3‴)建立了纯量和算子之间乘法的补充定律.

定义1 一个环K, 它同时又是域$\mathfrak{K}$上的向量空间, 且对任意$\lambda\in\mathfrak{K}$, $a, b\in K$有$\lambda(ab)=(\lambda a)b=a(\lambda b)$, 就称为是$\mathfrak{K}$上的**代数**. K作为向量空间的维数即称为代数K的**维数**, 所有的对K的乘法封闭的($L\cdot L\subseteq L$)的子空间$L\subset K$都称为是K的**子代数**.

说到代数, 大多指结合代数$(ab)c=a(bc)$且有$1:1\cdot x=x, x\in K$. 向量空间V的线性算子代数$\mathcal{L}(V)$就是这样. 代数$\mathcal{L}(V)$的矩阵变形$M_n(\mathfrak{K})$已经在[BA I]第2章§3中遇到过. 在那里对矩阵导出了类似的(3′)—(3‴). 尽管包括§1在内已经不止一次地提到线性映射和矩阵之间的对应, 我们还是应该再次把简单而重要的事实铭刻在脑海里: 如果

$$\mathcal{A}:\mathbf{e}_k\mapsto\mathcal{A}\mathbf{e}_k=\sum_{i=1}^{n}a_{ik}\mathbf{e}_i,\quad \mathcal{B}:\mathbf{e}_j\mapsto\mathcal{B}\mathbf{e}_j=\sum_{k=1}^{n}b_{kj}\mathbf{e}_k$$

是线性空间V在基底$(\mathbf{e}_i)$之下以$A=(a_{ij}), B=(b_{kj})$为矩阵的线性算子, 那么, 算子$\mathcal{A}\mathcal{B}$在同一个基底之下的矩阵是$C=AB$.

事实上

$$\begin{aligned}
\sum_i c_{ij}\mathbf{e}_i&=(\mathcal{A}\mathcal{B})\mathbf{e}_j=\mathcal{A}(\mathcal{B}\mathbf{e}_j)=\mathcal{A}\left(\sum_k b_{kj}\mathbf{e}_k\right)=\sum_k b_{kj}\mathcal{A}\mathbf{e}_k\\
&=\sum_k b_{kj}\sum_i a_{ik}\mathbf{e}_i=\sum_i\left(\sum_k a_{ik}b_{kj}\right)\mathbf{e}_i,
\end{aligned}$$

也就是$c_{ij}=\sum_k a_{ik}b_{kj}$和$(c_{ij})=AB$.

对于附加各种条件的环最有意义的是代数. 多项式代数$\mathfrak{K}[t]$是一个最简单的无穷维代数的例子. 在n^2维结合代数$\mathcal{L}(V)$中(见(2), $n=\dim V$)特别值得提到的是由单独一个算子生成的子代数. 即, 如果$\mathcal{A}$是个线性算子, 那么, 由它生成的子代数$\mathfrak{K}[\mathcal{A}]$是包

含$\mathcal{A}$的最小的子代数. 在这种特殊情形要充分认清, 单位算子$\mathcal{E}$属于$\mathfrak{K}[\mathcal{A}]$.子代数$\mathfrak{K}[\mathcal{A}]$的元素是算子$\mathcal{A}$的所有的各种可能的幂次

$$\mathcal{A}^0=\mathcal{E},\mathcal{A},\mathcal{A}^2=\mathcal{A}\mathcal{A},\cdots,\mathcal{A}^k=\underbrace{\mathcal{A}\mathcal{A}\cdots\mathcal{A}}_{k},\cdots,$$

以及它们的线性组合. 换句话来说, 如果

$$f(t)=a_0t^m+a_1t^{m-1}+\cdots+a_{m-1}t+a_m\in\mathfrak{K}[t],$$

那么

$$f(\mathcal{A})=a_0\mathcal{A}^{m-1}+\cdots+a_{m-1}\mathcal{A}+a_m\mathcal{E} \tag{4}$$

就是$\mathfrak{K}[\mathcal{A}]$中的线性算子的最一般的形式. 放到函数的观点之下, 我们就是说, $f(\mathcal{A})$是多项式$f\in\mathfrak{K}[t]$在$t=\mathcal{A}$时的值. 形如(4)的线性算子$f(\mathcal{A})$用自然的方式作用在$\mathbf{x}\in V$上:

$$f(\mathcal{A})\mathbf{x}=a_0\mathcal{A}^m\mathbf{x}+a_1\mathcal{A}^{m-1}\mathbf{x}+\cdots+a_{m-1}\mathcal{A}\mathbf{x}+a_m\mathbf{x}.$$

代数$\mathfrak{K}[\mathcal{A}]$是交换的, 因为

$$f(\mathcal{A})\cdot g(\mathcal{A})=g(\mathcal{A})\cdot f(\mathcal{A})$$

(方幂置换: $\mathcal{A}^k\cdot\mathcal{A}^l=\mathcal{A}^{k+l}=\mathcal{A}^l\cdot\mathcal{A}^k$的推论), 它的维数是多少呢? 后面我们会看到, 总有

$$\dim\mathfrak{K}[\mathcal{A}]\leqslant\dim V. \tag{5}$$

但这是个比较精细的结果, 暂且我们只能得到初步有用的说明.

定义2 称多项式$f(t)$**零化**线性算子$\mathcal{A}$, 如果$f(\mathcal{A})=\mathcal{O}$. 零化$\mathcal{A}$的次数最低的(首项系数为1的)多项式称为是算子$\mathcal{A}$的**极小多项式**.

设

$$\mu_{\mathcal{A}}(t)=t^m+\mu_1t^{m-1}+\cdots+\mu_{m-1}t+\mu_m \tag{6}$$

是算子$\mathcal{A}$的极小多项式. 那么, $\mathcal{E},\mathcal{A},\cdots,\mathcal{A}^{m-1}$必然线性无关. 因为, 若有$\sum\limits_{i=0}^{m-1}\lambda_i\mathcal{A}_i=\mathcal{O}$, 就意味着多项式$\sum\limits_{i=0}^{m-1}\lambda_it^i$零化$\mathcal{A}$, 然而它的次数却小于$m$. 反之, 若$\mathcal{E},\mathcal{A},\cdots,\mathcal{A}^{m-1}$(作为$\mathcal{L}(V)$的向量)线性无关, 而$\mathcal{A}^m$已经由它们线性表示出来, 那么就意味着$m$是极小多项式的次数. m的存在性是结论$\mathfrak{K}[\mathcal{A}]\subset\mathcal{L}(V)$的一个平凡的推论. 因为$\dim\mathcal{L}(V)=n^2$, 所以, $m\leqslant n^2$. 这就部分地证明了下面的命题.

定理1 所有线性算子$\mathcal{A}$都有极小多项式$\mu_{\mathcal{A}}(t)$. 它的次数与代数$\mathfrak{K}[\mathcal{A}]$的维数一致. 算子$\mathcal{A}$可逆, 当且仅当多项式(6)中的自由项$\mu_m$不等于零.

证明 定理的结论部分恰如我们上面引进的第一部分证明那样简单. 进而, 如果$\mu_m\neq0$, 那么

$$\mathcal{O}=\mu_{\mathcal{A}}(\mathcal{A})=\mathcal{A}(\mathcal{A}^{m-1}+\mu_1\mathcal{A}^{m-2}+\cdots+\mu_{m-1}\mathcal{E}).$$

这意味着, $\mathcal{A}$有零因子$\mathcal{A}^{m-1}+\mu_1\mathcal{A}^{m-2}+\cdots+\mu_{m-1}\mathcal{E}\neq\mathcal{O}$ ($\mu_{\mathcal{A}}(t)$的极小性), 而环中的零因子不能有逆元. 反过来, 如果$\mu_m\neq 0$, 那么, 关系式

$$\mathcal{A}(-\mu_m^{-1}\mathcal{A}^{m-1}-\mu_m^{-1}\mu_1\mathcal{A}^{m-2}-\cdots-\mu_m^{-1}\mu_{m-1}\mathcal{E})=\mathcal{E},$$

就在一个显式中给出了$\mathcal{A}$的逆算子. □

定理2 任意一个能零化算子$\mathcal{A}$的多项式$f(t)$用多项式$\mu_{\mathcal{A}}(t)$去除均无剩余项.

证明 据假设, 线性算子$f(\mathcal{A})$ (见(4))等于算子$\mathcal{O}$. 如果, $f(t)=q(t)\mu_{\mathcal{A}}(t)+r(t)$是$f(t)$被$\mu_{\mathcal{A}}(t)$除得到的剩余$r(t)$的结果, 那么

$$\mathcal{O}=f(\mathcal{A})=q(\mathcal{A})\cdot\mathcal{O}+r(\mathcal{A}),$$

从而$r(\mathcal{A})=\mathcal{O}$. 但$\deg r(t)<\deg\mu_{\mathcal{A}}(t)$, 这与极小多项式的定义相矛盾, 有$r(t)=0$.

□

定义3 线性算子$\mathcal{A}$称为是**幂零**的, 如果对某个$m>0$有$\mathcal{A}^m=\mathcal{O}$; 这类自然数$m$中的最小者称之为$\mathcal{A}$的**幂零指数**.

显然, 对于幂零指数为m的算子$\mathcal{A}$, 有$\mu_{\mathcal{A}}(t)=t^m$, 而对于满足$\mathcal{A}^2=\mathcal{A}$的非平凡算子$\mathcal{A}$, 有$\mu_{\mathcal{A}}(t)=t^2-t$. 同样, $\mu_{\mathcal{O}}(t)=t$, $\mu_{\mathcal{E}}(t)=t-1$.作用在次数$\leqslant n-1$的多项式空间P_n上的微分算子$\mathcal{D}_t$, 可以作为幂零指数为n的线性算子的典型例子. 投影算子$\mathcal{P}$(第1目中的例4)具有性质$\mathcal{P}^2=\mathcal{P}$. 这些例子在将来都会频繁地得到运用.

3. 线性算子在不同基底之下的矩阵 设V是域$\mathfrak{K}$上的n维向量空间, $\mathcal{A}:V\to V$是一个线性算子. 在V中选择一个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$, 我们可以得到矩阵$A=(a_{ki})$使得

$$\mathcal{A}\mathbf{e}_i=\sum_k a_{ki}\mathbf{e}_k. \tag{7}$$

但是, 同样这个算子$\mathcal{A}$在空间V的另外一个基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$之下将有另外一个矩阵$A'=(a'_{kj})$使得

$$\mathcal{A}\mathbf{e}'_j=\sum_k a'_{kj}\mathbf{e}'_k. \tag{7'}$$

如果$B=(b_{ij})$是从基底$(\mathbf{e}_i)$向基底$(\mathbf{e}'_j)$的转换矩阵, 那么$\mathbf{e}'_j=\sum_i b_{ij}\mathbf{e}_i,\quad 1\leqslant j\leqslant n$.这就提供了引入算子$\mathcal{B}$:

$$\mathcal{B}\mathbf{e}_j=\mathbf{e}'_j \tag{8}$$

的想法, 它在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下对应矩阵$B$. 因为$\det B\neq 0$ (第1章§2的定理4), 所以算子$\mathcal{B}$是可逆的.

最后, 我们定义一个辅助算子$\mathcal{A}'$, 它在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下与算子$\mathcal{A}$在基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$之下有同样一个矩阵$A'$. 换句话说, 令

$$\mathcal{A}'\mathbf{e}_j=\sum_i a'_{ij}\mathbf{e}_i. \tag{9}$$

我们有权这么做, 因为在固定的基底之下, 线性算子与矩阵之间有双射对应. 用(7)和(8) 把关系式(7′)写成

$$\mathcal{A}\mathcal{B}\mathbf{e}_j = \mathcal{A}\mathbf{e}'_j = \sum_i a'_{ij}\mathbf{e}'_i = \sum_i a'_{ij}\mathcal{B}\mathbf{e}_i = \mathcal{B}\left(\sum_i a'_{ij}\mathbf{e}_i\right),$$

由$\mathcal{B}$的可逆性以及对于$\mathcal{A}'$的表达式(9), 就得到

$$\mathcal{B}^{-1}\mathcal{A}\mathcal{B}\mathbf{e}_j = \mathcal{A}'\mathbf{e}_j, \quad 1 \leqslant j \leqslant n. \tag{10}$$

研究在同一个基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的所有算子$\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}'$, 我们可以把(10)转化成矩阵关系式

$$A' = B^{-1}AB. \tag{11}$$

可以用更直接的坐标方法推导出关系式(11). 照惯例, 设$\sum\limits_i x_i\mathbf{e}_i = \mathbf{x} = \sum\limits_i x'_i\mathbf{e}'_i$记录任意向量$\mathbf{x} \in V$向新的(打撇的)基底的转变, $X = [x_1, \cdots, x_n]$, $X' = [x'_1, \cdots, x'_n]$是对应的坐标列. 其次, 令$Y = AX, Y' = A'X'$, 其中$A$, A'是由关系式(7), (7′)决定的矩阵. 因为$X = BX', Y = BY'$ (见第1章§2的(4′)), 所以

$$ABX' = AX = Y = BY' = BA'X'.$$

由于列X' (向量$\mathbf{x} \in V$) 选择的任意性, 我们就有$AB = BA'$, 从而有$A' = B^{-1}AB$.

于是, 我们再次深信下面定理的正确性.

定理3 线性算子$\mathcal{A}$在基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$之下的矩阵$A'$可以用同一算子$\mathcal{A}$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下对应的矩阵$A$按照公式(11)得出, 其中$B$是由$(\mathbf{e}_i)$向$(\mathbf{e}'_j)$的转换矩阵.

定义4 称矩阵A'**相似**于矩阵A, 并记成$A' \sim A$, 如果存在非退化矩阵B按照(11) 把A和A'联系起来. 这里预先假定, 所有的矩阵都是同一个阶数的系数取自同一域$\mathfrak{K}$的方阵.

显然, 恒有$A \sim A$ (用$B = E$联系). 其次, 关系式(11)也可写成$A = B_1^{-1}A'B_1$, 其中$B_1 = B^{-1}$, 这表明, 相似关系是对称的: $A' \sim A \Rightarrow A \sim A'$.它同样是传递的: 如果$A' = B^{-1}AB, A'' = C^{-1}A'C$, 那么$A'' = (BC)^{-1}A(BC)$.即$A'' \sim A', A' \sim A \Rightarrow A'' \sim A$.这就是说, 相似关系是个等价关系, 进而, 所有的n阶方阵就被分成不相交的矩阵类(与[BA Ⅰ]第2章§3第6目的等价矩阵的等价类相比较). 按照定理3, 每个线性算子同样都对应一个相似矩阵类. 而相似的矩阵可以充当同一个线性算子在不同基底之下的矩阵.

线性算子语言在理论研究上很方便, 但大量的具体的计算还得以矩阵形式实现. 所以, 精确到相似的矩阵分类从实践的观点看是非常重要. 比方说, 如果我们需要对大的指数$k \geqslant 1000$(在实际中常常遇到这样的课题)计算一个$n > 1$阶矩阵(或者, 甚

至$n \geqslant 100)$$A$的方幂$A^k$, 那么, 自然地尝试去找一个容易计算其方幂$(A')^k$的$A' \sim A$. 如果有幸能得到一个矩阵$A'=\text{diag}\ (\lambda_1, \cdots, \lambda_n)$刚好与在$A$代表的相似类中, 那么, $A = BA'B^{-1}$, 从而$A^k = B\text{diag}\left(\lambda_1^k, \cdots, \lambda_n^k\right) B^{-1}$.在这种对应中, 矩阵$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$相当典型, 它直接对应的斐波那契数也是[BA Ⅰ]第2章§3第5目例3中的主要角色. 我们还要再多说一点, 能计算出来幂A^k就给出了对任意一个多项式$f(t) = a_0t^m + \cdots + a_m$找出值

$$f(A) = a_0A^m + a_1A^{m-1} + \cdots + a_{m-1}A + a_mE$$

的可能性.

4. 线性算子的行列式与迹　设$\mathcal{A}$是V上的一个线性算子. $\mathcal{A}$在V的某一个基底之下对应矩阵A, 则把A的行列式$\det\ A$称为线性算子$\mathcal{A}$的**行列式**. 因为$\det(B^{-1}AB) = \det A$, 所以$\det \mathcal{A}$是$\mathcal{A}$的一个不变量. 可逆矩阵对应可逆算子, 所以, $\det\ \mathcal{A} \neq 0$是算子$\mathcal{A}$可逆的一个充要条件. 在$\det\ \mathcal{A} = 0$的情形, 我们就要处理退化的线性算子$\mathcal{A}$.

现在, 称表达式

$$\text{tr}\ \mathcal{A} = \text{tr}\ A = \sum_{i=1}^{n} a_{ii}$$

为线性算子$\mathcal{A}$的**迹**, 其中$A = (a_{ij})$是算子$\mathcal{A}$对应的矩阵(tr是英文trace的缩写). 正如所知, 也容易证明, 对任意同一阶数的矩阵A, B, 有

$$\text{tr}\ AB = \text{tr}\ BA \tag{12}$$

把这个关系式用到$B^{-1}A$和B, 其中B是非退化的, 就有

$$\text{tr}(B^{-1}AB) = \text{tr}(B \cdot B^{-1}A) = \text{tr}\ A.$$

这意味着线性算子的迹的定义是适定的, 即和V的基底的选择没有关系. 与(12)类似的是

$$\text{tr}\ \mathcal{AB} = \text{tr}\ \mathcal{BA}. \tag{12'}$$

两个要引进的函数det和tr: $\mathcal{L}(V) \to \mathfrak{K}$有重要作用. 函数det是可乘的($\det \mathcal{AB} = (\det \mathcal{A})(\det \mathcal{B})$), 而且, 借助于它可引导出空间$V$的自同构群Aut V, 或者, 等价于V上所有非退化线性算子作成的群. 容易猜想到, 在V的任意选定的基底之下, 群Aut V就转化为在[BA Ⅰ]中已知的$n = \dim_{\mathfrak{K}} V$阶的完全线性群$GL_n(\mathfrak{K})$. 更准确些: 有群的同构Aut $V \cong GL_n(\mathfrak{K})$.

函数tr是线性的

$$\text{tr}(\alpha\mathcal{A} + \beta\mathcal{B}) = \alpha\text{tr}\ \mathcal{A} + \beta\text{tr}\ \mathcal{B}$$

(容易验证), 而这个事实被广泛地应用在数学领域, 比如, 内容丰富的群的特征标理论(见[BA Ⅲ])就整个地建立在迹的概念上. 看一个更“朴实”的应用.

例6 (**李**(Lie)**代数**) 一个代数, 作为环, 并不一定要具有结合律. 非结合环的最重要的例子就是所谓**李代数**(或**Lie代数**, 表示尊重S. Lie (1842—1899)), 在这个代数里, 乘法运算$(\mathbf{x},\mathbf{y}) \mapsto \mathbf{x}*\mathbf{y}$满足两条公理

i) $\mathbf{x}*\mathbf{x}=\mathbf{0}$; 进而$(\mathbf{x}+\mathbf{y})*(\mathbf{x}+\mathbf{y})=\mathbf{0} \Rightarrow \mathbf{x}*\mathbf{y}=-\mathbf{y}*\mathbf{x}$ (反对称性);

ii) $(\mathbf{x}*\mathbf{y})*\mathbf{z}+(\mathbf{y}*\mathbf{z})*\mathbf{x}+(\mathbf{z}*\mathbf{x})*\mathbf{y}=\mathbf{0}$(雅可比恒等式).

运算$\mathbf{x}*\mathbf{y}$很多时候被表成[x,y], 而且称为**换位子运算**. 向量空间$L=\mathbb{R}^3=\langle \mathbf{e}_i\rangle$是个以下面的向量的(外部的)向量乘积为其乘法运算的3维的李代数: 如果,

$$\mathbf{x}=x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3, \quad \mathbf{y}=y_1\mathbf{e}_1+y_2\mathbf{e}_2+y_3\mathbf{e}_3$$

那么,

$$[\mathbf{x},\mathbf{y}]=(x_2y_3-x_3y_2)\mathbf{e}_1+(x_3y_1-x_1y_3)\mathbf{e}_2+(x_1y_2-x_2y_1)\mathbf{e}_3.$$

其次, 容易验证(这样的验证对于每个搞数学的人的生活中都应当有), 在$L=\mathcal{L}(V)$上按规则

$$[\mathcal{A},\mathcal{B}]=\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A} \tag{13}$$

给出的新的乘法运算能使公理i), ii)得到满足, 从而可以把$\mathcal{L}(V)$看成是一个李代数. 通常用符号$\mathfrak{gl}_n(\mathfrak{K})$来代表它. 还有一种深刻的理论, 据这种理论, $\mathfrak{K}$上每一个有限维李代数都是一个李代数($\mathcal{L}(V)$; [,])的子代数, 其中V是$\mathfrak{K}$上的一个有限维向量空间(回忆一下, 子代数就是$\mathcal{L}(V)$的在乘法[,]之下封闭的子空间).

有限维与无穷维李代数在量子力学中发挥着非常本质性的作用(见补充文献中的教学参考书[2]). 事实在于, 量子理论中被称为动态变量者服从代数中的非交换定律, 而且它的非交换的程度, 计算起来, 恰恰就是“换位子”(13). 如果我们把V取成$\mathfrak{K}$上所有多项式的无穷维向量空间, 就得到一个平凡的, 但在某种狭窄意义下接近量子理论中交换关系的例子. 设$\mathcal{D}_t=d/dt$是对t的微分算子, 而$\mathcal{F}_t$是用t乘导致的乘法算子: $\mathcal{D}_t(f)=f', \mathcal{F}_t(f)=t\cdot f$, 容易验证

$$[\mathcal{D}_t,\mathcal{F}_t]=\mathcal{D}_t\mathcal{F}_t-\mathcal{F}_t\mathcal{D}_t=\mathcal{E} \tag{14}$$

是$V=\mathfrak{K}[t]$上的恒等算子.

问题来了: 形如(14)的关系式$[\mathcal{A},\mathcal{B}]=\mathcal{E}$能够在有限维代数$\mathcal{L}(V)$中实现吗? 这个问题的回答原来与基础域的特征数有关系. 如果$\mathfrak{K}$=$\mathbb{C}$和$\mathfrak{K}$=$\mathbb{R}$ (这是最有意义的情形), 那么, 立刻可导出矛盾

$$0=\operatorname{tr}\,\mathcal{A}\mathcal{B}-\operatorname{tr}\,\mathcal{B}\mathcal{A}=\operatorname{tr}[\mathcal{A},\mathcal{B}]=\operatorname{tr}\mathcal{E}=n=\dim V.$$

但是在$p|n$的情形, 其中$p=\operatorname{char}\mathfrak{K}$, 正如下面$\mathfrak{K}$上的$n$阶矩阵

$$J_p=\begin{pmatrix}0\,1\,0\cdots 0\,0\\ 0\,0\,1\cdots 0\,0\\ \cdots\cdots\cdots\\ 0\,0\,0\cdots 1\,0\\ 0\,0\,0\cdots 0\,1\\ 0\,0\,0\cdots 0\,0\end{pmatrix},N_p=\begin{pmatrix}0\,0\cdots & 0 & 0\\ 1\,0\cdots & 0 & 0\\ 0\,2\cdots & 0 & 0\\ \cdots\cdots & & \\ 0\,0\cdots & p-1 & 0\end{pmatrix} \tag{15}$$

运算表明的那样, 这个矛盾可以被消除. 前面利用函数tr的判别方法在这里不再起作用, 而事实上$[J_p,N_p]=E_p$.

习　　题

1. 验证, 形如(15)的两个矩阵都是幂零的: $J_p^p=N_p^p=0$.

2. 证明, 如果A,B,C分别是$n\times p,p\times q,q\times n$的矩阵, 那么, $\operatorname{tr}(ABC)=\operatorname{tr}(BCA)=\operatorname{tr}(CAB)$.

3. 把$GL_n(\mathbb{F}_p)$理解为p元域$\mathbb{F}_p$上的n维向量空间V的自同构群, 求出阶数$|GL_n(\mathbb{F}_p)|$.

4. 指明, 所有迹为零的线性算子的集合$\mathfrak{sl}_n(\mathfrak{K})$是李代数$\mathfrak{gl}_n(\mathfrak{K})=\mathcal{L}(V)$的一个余维数为1的子代数.

5. 证明, 对V上的任意线性算子$\mathcal{A}$, $\mathcal{B}$有等式

$$\operatorname{rank}\mathcal{A}=\operatorname{rank}\mathcal{B}+\dim(\operatorname{Im}\mathcal{A}\cap\operatorname{Ker}\mathcal{B}).$$

6. 利用习题5, 证明, 对于V上的任意线性算子$\mathcal{A},\mathcal{B},\mathcal{C}$而言, **弗罗贝尼乌斯不等式**

$$\operatorname{rank}\mathcal{BA}+\operatorname{rank}\mathcal{AC}\leqslant\operatorname{rank}\mathcal{A}+\operatorname{rank}\mathcal{BAC}$$

成立.

7. 证明, 对任意线性算子$\mathcal{A}$: $V\to V$和任意$i\geqslant 1$都有公式

$$\dim(\operatorname{Im}\ \mathcal{A}^{i-1}\cap\operatorname{Ker}\mathcal{A})=\dim\operatorname{Ker}\mathcal{A}^i-\dim\operatorname{Ker}\mathcal{A}^{i-1}$$

(对于i=1, 公式是显然的, 只要考虑到总有$\mathcal{A}^0=\mathcal{E}$的定义).

8. 证明, 两个矩阵$A,B\in M_n(\mathbb{R})$, 如果它们在复数域上相似, 那么, 就一定在实数域上相似.

9. 类似于定义2, 称$f(t)$是线性算子$\mathcal{A}$对于一个向量$\mathbf{v}\in V$的**零化多项式**, 如果$f(\mathcal{A})\mathbf{v}=\mathbf{0}$. 对于向量$\mathbf{v}$的算子$\mathcal{A}$的零化多项中次数最低者被称为$\mathcal{A}$的对向量$\mathbf{v}$的**极小多项式**, 将其用$\mu_{\mathcal{A},\mathbf{v}}(t)$代表. 我们假定域$\mathfrak{K}$是无穷的. 证明:

(1) $\mu_{\mathcal{A},\mathbf{v}}(t)$能整除$\mu_{\mathcal{A}}(t)$;

(2) 存在$\mathbf{a}\in V$使$\mu_{\mathcal{A},\mathbf{a}}(t)=\mu_{\mathcal{A}}(t)$.

10. 设V是个向量空间, U,W是它的两个子空间, 而且

$$V=V_1\oplus V_2,\quad W=W_1\oplus W_2$$

是直和分解, 其中$W_i\subseteq V_i,i=1,2$.其次, 再设$\mathcal{P}_i$是V_i的平行于$V_j,j\neq i$的投影.

证明:

(1) 如果

$$V_1 = W_1 + U \cap V_1, \quad V_2 = W_2 + \mathcal{P}_2(U), \tag{*}$$

那么, $V = W + U$;

(2) 如果$V = W + U$且$\mathcal{P}_2(U) \cap W_2 = \mathbf{0}$, 那么, 对于$V_1$和$V_2$的分解式(*)是真的, 而且$W \cap U = W_1 \cap U$.

11. 证明, 系数在特征为零的域$\mathfrak{K}$上的矩阵$A \in M_n(\mathfrak{K})$, 它迹为零必相似于沿主对角线取零值的矩阵A' ($A' = (a'_{ij}), a'_{11} = a'_{22} = \cdots = a'_{nn} = 0$).

12. 习题11中$\mathrm{Char}\ \mathfrak{K} = 0$是个限制吗?

§3 不变子空间与特征向量

1. 投影 §2第1目的例4建立了V到两个子空间的直和分解与我们已经熟知的, 有性质$\mathcal{P}^2 = \mathcal{P}$的投影算子$\mathcal{P}$之间的关系. 反过来, 所有具有这种性质的算子都是投影算子. 我们在更一般的背景下证明这个论断.

设$V = W_1 \oplus W_2 \oplus \cdots \oplus W_m$是$m$个子空间的直和分解(见第1章§2第5目), 那么, 每个向量$\mathbf{x} \in V$可唯一地表为

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_m, \quad \mathbf{x}_i \in W_i$$

的形式, 而映射$\mathcal{P}_i : \mathbf{x} \mapsto \mathbf{x}_i$是$V$上的线性算子. 此外, 还有

$$\mathcal{P}_1 + \mathcal{P}_2 + \cdots + \mathcal{P}_m = \mathcal{E},$$

而且当$i \neq j$时, $\mathcal{P}_i\mathcal{P}_j = \mathcal{O}, \mathcal{P}_i^2 = \mathcal{P}_i$.最后

$$W_i = \mathcal{P}_i V = \{\mathbf{x} \in V | \mathcal{P}_i\mathbf{x} = \mathbf{x}\},$$
$$K_i = \mathrm{Ker}\ \mathcal{P}_i = W_1 + \cdots + \hat{W}_i + \cdots + W_m,$$

而且$\mathcal{P}_i$就是V在W_i上沿着K_i的投影.

定理1 设$\mathcal{P}_1, \cdots, \mathcal{P}_m : V \to V$是满足下列条件的有限个线性映射的集合

$$\sum_{i=1}^{m} \mathcal{P}_i = \mathcal{E}; \quad \mathcal{P}_i^2 = \mathcal{P}_i, \quad 1 \leqslant i \leqslant m; \quad \mathcal{P}_i\mathcal{P}_j = \mathcal{O}, i \neq j. \tag{1}$$

那么

$$V = W_1 \oplus \cdots \oplus W_m,$$

其中$W_i = \mathrm{Im}\ \mathcal{P}_i$.

证明 按条件, 对任意$\mathbf{x} \in V$, 我们有

$$\mathbf{x} = \mathcal{E}\mathbf{x} = \sum \mathcal{P}_i\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_m, \quad \mathbf{x}_i \in W_i.$$

这是因为$V = W_1 + \cdots + W_m$.这个和是直和, 利用第1章§2的第5目(定理7)的准则, 我们就可以确信这一点. 也就是说, 设$\mathbf{x} \in W_j \cap \sum_{i \neq j} W_i$, 因为$W_i = \operatorname{Im} \mathcal{P}_i$, 所以, 可找到那样的$\mathbf{x}_1, \cdots, \mathbf{x}_m$使得

$$\mathbf{x} = \mathcal{P}_j(\mathbf{x}_j) = \sum_{i \neq j} \mathcal{P}_i(\mathbf{x}_i).$$

把算子$\mathcal{P}_j$作用到这个等式上去, 并利用定义的性质$\mathcal{P}_j^2 = \mathcal{P}_j$; 当$i \neq j$时, $\mathcal{P}_j \mathcal{P}_i = \mathcal{O}$, 就得到

$$\mathbf{x} = \mathcal{P}_j(\mathbf{x}_j) = \mathcal{P}_j^2(\mathbf{x}_j) = \sum_{i \neq j} \mathcal{P}_j \mathcal{P}_i(\mathbf{x}_i) = \mathbf{0}.$$

这样一来, $V = \sum W_i$是个直和而$\mathcal{P}_i$就是V在W_i上沿着$K_i = \operatorname{Ker} \mathcal{P}_i = \sum_{j \neq i} W_j$的投影. □

我们再补充一点, 如果$\mathcal{P}^2 = \mathcal{P}$而$V = U \oplus W$是与这个投影算子相联系的直和分解且$U = \operatorname{Im} \mathcal{P} = \langle \mathbf{e}_1, \cdots, \mathbf{e}_r \rangle, W = \operatorname{Ker} \mathcal{P} = \langle \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle$, 那么, 在这个选定的基底之下算子$\mathcal{P}$对应矩阵

$$P = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}, r = \operatorname{rank} \mathcal{P}. \tag{2}$$

特别地, 可以看出, 具有性质$A^2 = A$的秩为r的n阶矩阵A必相似于矩阵$P: B^{-1}AB = P$, $\operatorname{rank} A = \operatorname{tr} A$.

说明 经常说, 满足关系式

$$\mathcal{P}_i \mathcal{P}_j = \delta_{ij} \mathcal{P}_i, \quad 1 \leqslant i, j \leqslant m,$$

的算子$\mathcal{P}_1, \cdots, \mathcal{P}_m$组成一个**正交的等方算子组**$\{\mathcal{P}_i | 1 \leqslant i \leqslant m\}$, 而它们对应的矩阵组成**正交的等方矩阵组**$\{P_i | 1 \leqslant i \leqslant m\}$. 如果满足(1)的所有条件, 就说成是**完全正交组**.

2. 不变子空间 每个线性算子$\mathcal{A}: V \to V$都不仅仅能作用在一个一个的向量$\mathbf{x} \in V$上, 而且可以作用在子空间$U \subset V$上: $\mathcal{A}U = \{\mathcal{A}\mathbf{x} | \mathbf{x} \in U\}$. 与这个重要的有特殊价值的思想相联系可得出不变性的概念

定义1 子空间$U \subset V$相对于线性算子$\mathcal{A}: V \to V$是**不变的**, 如果$\mathcal{A}U \subset U$.

例如, $\operatorname{Ker} \mathcal{A}$和$\operatorname{Im} \mathcal{A}$都是$\mathcal{A}$的不变子空间, 尽管它们可能是平凡的, 即与$\{\mathbf{0}\}$或$V$重复. 对于次数$\leqslant n-1$的多项式空间$P_n$上的微分算子$\mathcal{D}_t$立刻可以做出一个次数小于$i-1$ 的多项式的不变子空间$V_i (i = 1, 2, \cdots, n)$的链

$$\{0\} \subset V_1 \subset V_2 \subset \cdots \subset V_n = V \tag{3}$$

上面(第1目)看到的投影算子组$\mathcal{P}_1, \cdots, \mathcal{P}_m$在一些关系式中是很有意义的, 这些关系式可联想到大量的对每个$\mathcal{P}_1, \cdots, \mathcal{P}_m$都不变的子空间

$$W_{i_1} \oplus W_{i_2} \oplus \cdots \oplus W_{i_k}, \quad \{i_1, \cdots, i_k\} \subseteq \{1, 2, \cdots, m\},$$

(我们这里用到了一个显而易见的事实, 在V中对线性算子$\mathcal{A}$不变的子空间的和与交也总是不变的).

在§2末尾出示的矩阵J_p, N_p可以充当与之对立的例子. 它们对应的算子作用在p维空间$V = \mathfrak{K}^p$ (在特征数为$p > 0$的域$\mathfrak{K}$上)上, 如此一来, 在这里没有不变子空间. 可以指出这种差别的一个根本原因: $[\mathcal{P}_i, \mathcal{P}_j] = \mathcal{O}$而$[J_p, N_p] \neq 0$.

在域$\mathbb{R}$上同样可能出现类似的现象. 算子$\mathcal{A}$把$\mathbb{R}^2$旋转α角($0 < \alpha < \pi$; 见§2第1目之例3)就没有非平凡的不变子空间, 它应该是在$\mathcal{A}$作用下自己变成自己的一条直线才行.

固有的不变子空间$\{\mathbf{0}\} \subset U \subset V$的存在提供了一种通过选择$V$中的适当的基底简化算子$\mathcal{A}$的矩阵$A$的可能性. 也就是说, 如果把$U$的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m)$的基底扩充成$V$的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m, \mathbf{e}_{m+1}, \cdots, \mathbf{e}_n)$. 那么, 由条件$\mathcal{A}\mathbf{e}_i \in U, \quad 1 \leqslant i \leqslant m$, 得到在这个基底之下算子$\mathcal{A}$的矩阵就是

$$A = \begin{pmatrix} A_1 & A_0 \\ 0 & A_2 \end{pmatrix}, \tag{4}$$

其中, A_1是$m \times m$阶矩阵, A_2是$(n-m) \times (n-m)$阶矩阵, 而A_0是$m \times (n-m)$阶矩阵. 可以把A_1看成是算子$\mathcal{A}$限制在U上的算子$\mathcal{A}_U$的矩阵(设$A_1 = A_U$是方便的).

我们随时可以设想A_0是零矩阵. 此时, 显然, $W = \langle \mathbf{e}_{m+1}, \cdots, \mathbf{e}_n \rangle$也是个$V$的不变子空间, 而$A_2$就是算子$\mathcal{A}_W$的矩阵, 这种情形说成是$V$的不变子空间的直和$V = U \oplus W$的分解式对应的算子的直和:

$$\mathcal{A} = \mathcal{A}_U \dot{+} \mathcal{A}_W. \tag{5}$$

算子的直和的矩阵具有分块对角形式

$$A = \begin{pmatrix} A_U & 0 \\ 0 & A_W \end{pmatrix} = A_U \dot{+} A_W. \tag{5$'$}$$

实际上, 我们证明了

定理2 *空间V是它的两个对于线性算子$\mathcal{A}: V \to V$不变的子空间U和W的直和, 当且仅当, 这个算子在某个基底之下取(5′)形式.*

这个论断可用极明显的方式照搬到任意m个不变子空间, 它们的直和等于V的情形. 当$m = n = \dim V$时, 我们就得到了线性算子的矩阵在适当基底之下化成对角形的条件.

我们还是返回到等式(4)中矩阵A_0不一定为零的情形, 尽管已经选择了向量$\mathbf{e}_{m+1}, \cdots, \mathbf{e}_n$扩充了不变子空间$U \subset V$的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m)$. 这表明, 虽然按照第1章§2之定理9, 线性空间V有多种方式分解成直和$V = U \oplus W$, 却可能没有一个补子空间W是对$\mathcal{A}$不变的. 这种情形最好不过是以微分算子$\mathcal{D}_t$为例证: 如果V_i是(3)中不变子空间链中的一个, 且$V = V_i \oplus W_i$, 那么, 显然地, $\mathcal{D}_t(W_i) \cap V_i \neq \mathbf{0}$.

注意这样的结论, 如果$\mathcal{A}(U)\subseteq U, \mathcal{B}(U)\subseteq U$, 也就是说, U对$\mathcal{A}$和$\mathcal{B}$是共同的不变子空间, 那么$\alpha\mathcal{A}+\beta\mathcal{B}$和乘积$\mathcal{AB}$, $\mathcal{BA}$都是不变的. 特别地,

$$\mathcal{A}U\subseteq U\Rightarrow f(\mathcal{A})U\subseteq U$$

对任意多项式$f\in\mathfrak{K}[t]$都成立.

3. 特征向量, 特征多项式　1维的不变子空间需要特别加以关注.

定义2　对于$\mathcal{A}$不变的1维子空间中的任意一个非零向量都被称为是算子$\mathcal{A}$的一个**特征向量**. 如果$\mathbf{x}$是个特征向量:

$$\mathcal{A}\mathbf{x}=\lambda\mathbf{x},$$

那么, 称纯量$\lambda\in\mathfrak{K}$是算子$\mathcal{A}$的一个对应特征向量$\mathbf{x}$的**特征值**. 有时也说是**本征值, 固有值, 本征向量**或**固有向量**.

注意

$$\mathcal{A}\mathbf{x}=\lambda\mathbf{x}\Rightarrow\mathcal{A}^k\mathbf{x}=\lambda^k\mathbf{x},$$

从而有

$$f(\mathcal{A})\mathbf{x}=f(\lambda)\mathbf{x}, \tag{6}$$

对任意一个$f\in\mathfrak{K}[t]$都成立. 特别地,

$$f(\mathcal{A})=\mathcal{O}\Rightarrow f(\lambda)=0 \tag{7}$$

对算子$\mathcal{A}$的所有特征值λ都成立.

设

$$V^\lambda=\{\mathbf{v}\in V|\mathcal{A}\mathbf{v}=\lambda\mathbf{v}\}$$

是由$\mathbf{0}$和所有与特征值λ相伴随的特征向量组成的子空间.

定义3　明显地, 蕴涵式

$$\mathcal{A}\mathbf{x}=\lambda\mathbf{x}, \mathcal{A}\mathbf{y}=\lambda\mathbf{y}\Rightarrow\mathcal{A}(\alpha\mathbf{x}+\beta\mathbf{y})=\lambda(\alpha\mathbf{x}+\beta\mathbf{y})$$

给出了算子$\mathcal{A}$与λ相伴随的特征子空间V^λ的称谓的基础. 它的维数$\dim V^\lambda$称为特征值λ的**几何重数**.

特征向量的存在性条件显然可以描述成下面的形式

$$(\mathcal{A}-\lambda\mathcal{E})\mathbf{x}=\mathbf{0},\quad \mathbf{x}\neq\mathbf{0}, \tag{8}$$

也就是

$$\operatorname{Ker}(\mathcal{A}-\lambda\mathcal{E})\neq\mathbf{0}.$$

这意味着, 算子$\mathcal{A}-\lambda\mathcal{E}$是退化的:

$$\det(\mathcal{A}-\lambda\mathcal{E})=0. \tag{9}$$

如果在向量空间V的某个基底$(\mathbf{e}_i)$之下，算子$\mathcal{A}$的矩阵$A=(a_{ij})$，那么，算子$\mathcal{A}-\lambda\mathcal{E}$的矩阵就是$A-\lambda E$，条件(9)就转化成

$$\det(A-\lambda E)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}=0. \tag{9'}$$

设$\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n$是算子$\mathcal{A}$的属于特征值$\lambda$的一个特征向量. 在矩阵记法之下，等式(8)形如

$$\begin{aligned} &(a_{11}-\lambda)x_1+a_{12}x_2+\cdots+a_{1n}x_n=0,\\ &a_{21}x_1+(a_{22}-\lambda)x_2+\cdots+a_{2n}x_n=0,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &a_{n1}x_1+a_{n2}x_2+\cdots+(a_{nn}-\lambda)x_n=0. \end{aligned}$$

就得到了这个行列式(9′)为零的齐次线性方程组必有非平凡解. 正如我们在[BA Ⅰ]中就已经知道的，这个方程组的所有解的子空间与算子$\mathcal{A}$的特征子空间V^λ相重合. 维数$\dim V^\lambda$(或λ的几何重数)等于$n-r$，其中r是矩阵$A-\lambda E$的秩.

按[BA Ⅰ]第3章§1公式(3)将行列式$\det(tE-A)=(-1)^n\det(A-tE)$展开:

$$\det(tE-A)=\sum_{\pi\in S_n}\varepsilon_\pi(\delta_{1,\pi1}t-a_{1,\pi1})(\delta_{2,\pi2}t-a_{2,\pi2})\cdots(\delta_{n,\pi n}t-a_{n,\pi n}),$$

我们就得到一个以$\chi_i\in\mathfrak{K}$为系数带有未知变量t的n次标准多项式

$$\chi_A(t)=\det(tE-A)=t^n+\chi_1t^{n-1}+\cdots+\chi_{n-1}t+\chi_n. \tag{10}$$

定义4 称多项式(10)是矩阵A的**特征多项式**. 方程式$\chi_A(t)=0$被称为是**特征方程式**(有时说是矩阵A的**长期方程式**).

实际上，我们可以讨论线性算子$\mathcal{A}$的特征多项式而不涉及任何与它相伴的矩阵A. 因为借助§2的定理3有并不复杂的

定理3 相似矩阵的特征多项式相同.

证明 设$A'=C^{-1}AC$. 那么，$\det(tE-A')=\det(tC^{-1}EC-C^{-1}AC)=\det[C^{-1}(tE-A)C]=\det C^{-1}\det(tE-A)\det C=\det(tE-A)$. □

于是，我们可以规定

$$\chi_{\mathcal{A}}(t):=\chi_A(t).$$

确定的等式(10)表明，纯量$\lambda\in\mathfrak{K}$是线性算子$\mathcal{A}$的特征值，当且仅当$\chi_{\mathcal{A}}(\lambda)=0$，即$\lambda$是特征多项式的一个根. 如果特征多项式$\chi_{\mathcal{A}}(t)$在$\mathfrak{K}$中没有根，那么算子$\mathcal{A}$就没有特征向量. 所有的作用在复的向量空间上的线性算子一定有特征向量.

定义5 λ作为特征多项式$\chi_{\mathcal{A}}(t)$的根的重数被称为是线性算子$\mathcal{A}$的特征值λ的**代数重数**.

定理4 特征值λ的几何重数不超过它的代数重数.

证明 按定义, 几何重数是方程式$\mathcal{A}\mathbf{x} = \lambda\mathbf{x}$的解空间$V^\lambda$的维数$m$. 显然, V^λ是对$\mathcal{A}$不变的. 而且若$\mathcal{A}'$是$\mathcal{A}$在V^λ上的限制, 那么$\det(t\mathcal{E}'-\mathcal{A}') = (t-\lambda)^m$, 同时, $\chi_{\mathcal{A}}(t) = (t-\lambda)^m q(t)$, 其中$q(t)$是$\mathfrak{K}[t]$中的某个多项式, 设$\lambda$是多项式$q(t)$的$k$重根, $k \geqslant 0$. 在这种情况下, λ的代数重数就是$m+k$. □

4. 可对角化的判别准则 特征多项式$\chi_{\mathcal{A}}(t)$的根(也说是**特征根**)组成了一个产生关于算子$\mathcal{A}$的重要信息的集合, 但是, 根据明显的理由, 并非所有的特征根都是平等的.

定义6 线性算子$\mathcal{A}$的所有特征值的集合称为是这个算子的**谱**, 并且记为$\mathrm{Spec}(\mathcal{A})$ (用几何重数计算特征值的个数). 可以类似地讨论矩阵A的**谱**$\mathrm{Spec}\ A$. 如果对应的几何重数是1, 谱点就说成是**单的**, 如果所有的谱点都是单的, 就说谱是**单的**.

在代数封闭域的情形, 例如, $\mathfrak{K} = \mathbb{C}$, 特征根与谱点是一致的, 而在一般情形, 谱可能是空的, 例如实平面上旋转算子的谱.

引理1 属于不同特征值的特征向量必然线性无关. 和$\sum\limits_{\lambda\in \mathrm{Spec}\ \mathcal{A}} V^\lambda$是直和(一般来说, $\sum\limits_\lambda V^\lambda$不一定与$V$重合).

证明 设$\lambda_1,\cdots,\lambda_m$是一些不同的特征值, $V^{\lambda_1},\cdots,V^{\lambda_m}$是它们对应的特征子空间. 在每个$V^{\lambda_i}$中选取一个特征向量$\mathbf{e}_i$. 需要证明它们线性无关. 当$m=1$时, 命题真确. 对$m$用数学归纳法且设存在非平凡的线性关系式

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_m\mathbf{e}_m = \mathbf{0},$$

比方说, 其中$\alpha_1 \neq 0$, 我们把$\mathcal{A}$作用到这个等式的两侧, 因为$\mathcal{A}\mathbf{e}_i = \lambda_i\mathbf{e}_i$, 所以

$$\alpha_1\lambda_1\mathbf{e}_1 + \alpha_2\lambda_2\mathbf{e}_2 + \cdots + \alpha_m\lambda_m\mathbf{e}_m = \mathbf{0}.$$

用λ_m乘第一个关系式, 然后从第二个等式中将其减去, 就得到前边的$m-1$个向量的关系式

$$\alpha_1(\lambda_m - \lambda_1)\mathbf{e}_1 + \cdots + \alpha_{m-1}(\lambda_m - \lambda_{m-1})\mathbf{e}_{m-1} = \mathbf{0}.$$

由归纳法假定, 得$\alpha_i(\lambda_m - \lambda_i) = 0$, $i = 1,2,\cdots,m-1$.但

$$\alpha_1 \neq 0, \quad \lambda_m \neq \lambda_i, \quad i < m \Rightarrow \alpha_1(\lambda_m - \lambda_1) \neq 0.$$

得到的矛盾证实了我们的论断.

按照定义, 任意非零向量$\mathbf{e}_i \in V^{\lambda_i}$都是特征向量. 所以, 根据证明, $V^{\lambda_i} \bigcap \sum\limits_{j\neq i} V^{\lambda_j} = \mathbf{0}$. 这就意味着, 和$\sum\limits_i V^{\lambda_i}$是直的. □

定义7 称n维向量空间V上的线性算子$\mathcal{A}$是**可对角化的**, 如果有基底$(\mathbf{e}_i)$, 对于这个基底, 算子$\mathcal{A}$的矩阵取对角形式.

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

定理5 具有单谱的算子$\mathcal{A}$是可对角化的.

证明 定理的条件要求多项式$\chi_{\mathcal{A}}(t)$在基础域$\mathfrak{K}$上有$n = \dim V$个不同的根$\lambda_1, \cdots, \lambda_n$. 它们对应特征向量$\mathbf{e}_i$, $i = 1, \cdots, n$. 据引理1, 这些向量线性无关. 这意味着$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle$, 也就是$\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$, 所以, $A = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. □

算子谱的单性根本上只是它可对角化的一个充分条件, 例如, 等方算子(见(2))是可对角化的, 虽然, 当$n > 2$时它的谱不是单的, 这个事实的内在理由可由下面的定理部分地解释.

定理6 设$\mathcal{A}$是域$\mathfrak{K}$上有限维向量空间V上的一个线性算子, $\mathcal{A}$可对角化的充要条件是满足下面的两个条件:

i) 特征多项式$\chi_{\mathcal{A}}(t)$的所有根都在$\mathfrak{K}$中;

ii) 每个特征值λ的几何重数都与自己的代数重数一致.

证明 设条件i),ii)被满足. 如果$\lambda_1, \cdots, \lambda_m$是多项式$\chi_{\mathcal{A}}(t)$的不同的根, 而$k_1, \cdots, k_m$是它们的重数, 那么

$$\dim V^{\lambda_i} = k_i, \quad k_1 + k_2 + \cdots + k_m = n. \tag{11}$$

据引理1, 任意一组同时都不为零的向量$\mathbf{v}_i \in V^{\lambda_i}$一定是线性无关的, $i = 1, \cdots, m$. 因此, 有

$$V^{\lambda_i} \cap (V^{\lambda_1} + \cdots + \widehat{V^{\lambda_i}} + \cdots + V^{\lambda_m}) = \mathbf{0}. \tag{12}$$

这意味着(见第1章§2的定理7), 和$V^{\lambda_1} + \cdots + V^{\lambda_m}$是直和. 考虑到等式(11), 我们就得到

$$V = V^{\lambda_1} + \cdots + V^{\lambda_m}. \tag{13}$$

把V^{λ_i}的基底合并起来即得到V的一个基底, 它是由n个线性无关的$\mathcal{A}$的特征向量组成的**特征基底**. 它的存在性等价于$\mathcal{A}$的可对角化性.

反过来: 设算子$\mathcal{A}$可对角化. 再一次用$\lambda_1, \cdots, \lambda_m$代表它的不同的特征根, 并设$l_i = \dim\ V^{\lambda_i}$, $1 \leqslant i \leqslant m$. 条件(12)亦然被满足, 从而有由子空间$V^{\lambda_i}$的元素组合的特征向量基底, 故, $V^{\lambda_1}, \cdots, V^{\lambda_m}$生成$V$. 由此, 我们可以得到结论, 等式(13)成立. 相应于由V^{λ_i}的基底并起来的V的基底, 算子$\mathcal{A}$的矩阵必然是

$$A = \mathrm{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}_{l_1}; \cdots; \underbrace{\lambda_m, \cdots, \lambda_m}_{l_m}).$$

由等式

$$\chi_{\mathcal{A}}(t) = \chi_{A}(t) = \det(tE - A) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_m)^{l_m}$$

可以推出，多项式$\chi_{\mathcal{A}}(t)$的所有根都在$\mathfrak{K}$中，也就是满足条件i). 同时，对$i = 1, 2, \cdots, m$，整数l_i与根λ_i的代数重数k_i(见(11))一致. □

5. 不变子空间的存在性　所有的有关于算子的不变子空间, 特征值和特征向量的推断, 原则上都是在任意域上进行的. 但是, 这些研究对象的存在性却与基础域有直接的关系, 在更加重要的域$\mathbb{R}$和$\mathbb{C}$上的例子使我们深信这一点.

定理7　*所有复的(相对照地，实的)线性算子$\mathcal{A}$必有1维(相对照地，1维或2维)的不变子空间.*

证明　因为$\chi_{\mathcal{A}}(t)$在$\mathbb{C}$上至少有一个根, 那么, 已知的寻求特征向量的方法很容易就给出源于空间V的不变子空间.

在实数域$\mathbb{R}$的情形，我们看算子$\mathcal{A}$的极小多项式$\mu_{\mathcal{A}}(t)$(见§2定义2)，它的系数都在$\mathbb{R}$中, 如果$\mu_{\mathcal{A}}(t)$有一个实根α, 那么

$$\mu_{\mathcal{A}}(t) = (t - \alpha)g(t), \quad g(t) \in \mathbb{R}[t].$$

由$\mu_{\mathcal{A}}(t)$的极小性知$g(\mathcal{A}) \neq \mathcal{O}$, 所以, $\mathbf{v} = g(\mathcal{A})\mathbf{u} \neq \mathbf{0}$, 对某个向量$\mathbf{u} \in V$成立. 但

$$(\mathcal{A} - \alpha\mathcal{E})\mathbf{v} = (\mathcal{A} - \alpha\mathcal{E})g(\mathcal{A})\mathbf{u} = \mu_{\mathcal{A}}(\mathcal{A})\mathbf{u} = \mathbf{0},$$

从而$\mathcal{A}\mathbf{v} = \alpha\mathbf{v}$, 这说明, $\mathbf{v}$是一个特征向量.

现在设, $\mathcal{A}$没有特征向量, 那么, 据上面的证明, $\mu_{\mathcal{A}}(t)$就没有实根, 但按着关于实系数多项式的定理([BA Ⅰ]第6章§4, 定理1), 我们可以记成

$$\mu_{\mathcal{A}}(t) = (t^2 - \alpha t - \beta)h(t), \quad \alpha, \beta \in \mathbb{R}, \quad h(t) \in \mathbb{R}[t].$$

再次有对某个$\mathbf{u} \in V$使$\mathbf{v} = h(\mathcal{A})\mathbf{u} \neq \mathbf{0}$, 且

$$\mathcal{A}^2\mathbf{v} - \alpha\mathcal{A}\mathbf{v} - \beta\mathbf{v} = \mu_{\mathcal{A}}(\mathcal{A})\mathbf{u} = \mathbf{0}.$$

这就得到了，$\mathcal{A}^2\mathbf{v} = \alpha\mathcal{A}\mathbf{v} + \beta\mathbf{v}$，又因为$\mathcal{A}\mathbf{v} \neq \lambda\mathbf{v}$ (没有1维子空间)，所以，$L = \langle \mathbf{v}, \mathcal{A}\mathbf{v} \rangle$是个2维的不变子空间. □

6. 共轭线性算子　看一下算子概念与共轭空间概念有怎样的联系. 设V是域$\mathfrak{K}$上的向量空间, V^*是和它共轭的空间. $\mathcal{A}$是V上的一个线性算子. 对任意一个固定的元素$f \in V^*$, 映射$\mathbf{x} \mapsto (f, \mathcal{A}\mathbf{x}) := f(\mathcal{A}(\mathbf{x}))$(按第1章§3中第2目中的记法)又是$V^*$中的一个元素, 也就是说, 它是个线性函数:

$$(f, \mathcal{A}(\alpha\mathbf{x} + \beta\mathbf{y})) = (f, \alpha\mathcal{A}\mathbf{x} + \beta\mathcal{A}\mathbf{y}) = \alpha(f, \mathcal{A}\mathbf{x}) + \beta(f, \mathcal{A}\mathbf{y}).$$

这样, 把符号$\mathcal{A}^*f$当成V上的一个线性函数, 我们就可以令

$$(\mathcal{A}^*f, \mathbf{x}) := (f, \mathcal{A}\mathbf{x}), \tag{14}$$

对应$\mathcal{A}^*: f \to \mathcal{A}^* f$以$f$为变量就定义了$V^* \to V^*$的线性映射:

$$(\mathcal{A}^*(\alpha f + \beta g), \mathbf{x}) = (\alpha f + \beta g, \mathcal{A}\mathbf{x}) = \alpha(f, \mathcal{A}\mathbf{x}) + \beta(g, \mathcal{A}\mathbf{x})$$
$$= \alpha(\mathcal{A}^* f, \mathbf{x}) + \beta(\mathcal{A}^* g, \mathbf{x}) = (\alpha\mathcal{A}^* f + \beta\mathcal{A}^* g, \mathbf{x}),$$

所以$\mathcal{A}^* \in \mathcal{L}(V^*)$.

定义8 关系式(14)给出的V^*上的线性算子$\mathcal{A}^*$被称为是与$\mathcal{A} \in \mathcal{L}(V)$**共轭的**算子.

这样, 我们就得到了一个$\mathcal{L}(V) \to \mathcal{L}(V^*)$的映射, 即$*: \mathcal{A} \to \mathcal{A}^*$. 由定义可以直接导出它的下列性质:

$$\mathcal{O}_V^* = \mathcal{O}_{V^*}, \mathcal{E}_V^* = \mathcal{E}_{V^*}, (\alpha\mathcal{A})^* = \alpha\mathcal{A}^*,$$

$$(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, (\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*. \tag{15}$$

作为一个例子, (15)中的最后一个关系式可以这样证明

$$((\mathcal{A}\mathcal{B})^* f, \mathbf{x}) = (f, (\mathcal{A}\mathcal{B})\mathbf{x}) = (f, \mathcal{A}(\mathcal{B}\mathbf{x})) = (\mathcal{A}^* f, \mathcal{B}\mathbf{x}) = (\mathcal{B}^*\mathcal{A}^* f, \mathbf{x}).$$

为了给出算子$\mathcal{A}^*$的矩阵形式, 很自然地V和V^*的对偶基$(\mathbf{e}_i)$和(e^j). 如果$\mathcal{A}\mathbf{e}_j = \sum_{k=1}^{n} a_{kj}\mathbf{e}_k$, 那么

$$(e^i, \mathcal{A}\mathbf{e}_j) = \sum_{k=1}^{n} a_{kj}(e^i, \mathbf{e}_k) = \sum_{k=1}^{n} a_{kj}\delta_{ik} = a_{ij}.$$

接着, 设

$$\mathcal{A}^* e^i = \sum_{k=1}^{n} a_{ki}^* e^k,$$

就有$(\mathcal{A}^* e^i, \mathbf{e}_j) = \sum_{k=1}^{n} a_{ki}^*(e^k, \mathbf{e}_j) = a_{ji}^*$. 另一方面, 因为在关系式(14)中有$(\mathcal{A}^* e^i, \mathbf{e}_j) = (e^i, \mathcal{A}\mathbf{e}_j) = a_{ij}$, 所以$a_{ji}^* = a_{ij}$, 从而有

定理8 如果在空间V的基底$(\mathbf{e}_i)$之下线性算子$\mathcal{A}$对应矩阵$A = (a_{ij})$. 那么, $\mathcal{A}$的共轭算子在空间V^*的对偶基底(e^i)之下对应矩阵${}^tA$:$A^* = (a_{ij}^*) = {}^t A$. □

我们注意有限维向量空间的自反性, 它用自然映射(第1章§3 定理2)结论给出了把V^{**}和V等同起来的可能性, 同样地, 算子可表达成

$$\mathcal{A}^{**} = \mathcal{A}. \tag{16}$$

事实上, 由于自反性, V^*的任意线性函数可以把自己想象成在某个固定的$\mathbf{x} \in V$之下, $f \mapsto (f, \mathbf{x})$. 特别地, $(\mathcal{A}^* f, \mathbf{x}) = (f, \mathbf{y})$. 按定义$\mathbf{y} = \mathcal{A}^{**}\mathbf{x}$. 可见

$$(f, \mathcal{A}\mathbf{x}) = (\mathcal{A}^* f, \mathbf{x}) = (f, \mathcal{A}^{**}\mathbf{x}),$$

从这里可以导出关系式(16). 它表明, 具有性质(15)的映射$\mathcal{A} \mapsto \mathcal{A}^{**}$是个一一对应, 称它为$\mathcal{L}(V)$与$\mathcal{L}(V^*)$之间的**反同构**.

把对$(V,\mathcal{A})$和$(V^*,\mathcal{A}^*)$同时加以研究常常会推导出有实践意义的结果. 在这些内容丰富的例子中, 有一个就是下面论断的证明.

定理9　*V上所有复的线性算子$\mathcal{A}$必有不变的超平面.*

证明　设$\dim V=n$. 如同我们知道的那样, 对于V上任意线性函数$f\neq 0$, 必有$\dim \operatorname{Ker} f=n-1$. 现在, 比方说, 取$f$为$V^*$上的线性算子$\mathcal{A}^*$的一个特征向量, 按定理7, 它一定是存在的. 如果λ是它对应的特征值, 那么, 由定义等式(14), $\mathbf{x}\in \operatorname{Ker} f\Rightarrow 0=\lambda(f,\mathbf{x})=(\lambda f,\mathbf{x})=(\mathcal{A}^*f,\mathbf{x})=(f,\mathcal{A}\mathbf{x})\Rightarrow \mathcal{A}\mathbf{x}\in \operatorname{Ker} f$. 这就意味着, $\operatorname{Ker} f$就是所要求的超平面. □

7. 商算子　设L是作用在向量空间V上的线性算子$\mathcal{A}$的一个不变子空间. 把V和L看成是固定的, 并且约定把第1章§2第6目中定义的商空间V/L用符号$\overline{V}$记之, 再用$\overline{\mathbf{x}}$记为它的元素$\mathbf{x}+L$.

定义9　*在空间$\overline{V}$上, 用关系式$\overline{\mathcal{A}}\cdot\overline{\mathbf{x}}=\overline{\mathcal{A}\mathbf{x}}$导出一个***商算子***, 换言之,* $\overline{\mathcal{A}}(\mathbf{x}+L)=\mathcal{A}\mathbf{x}+L$.

这个定义与代表元$\mathbf{x}$的选择没有关系: 如果$\mathbf{x}+L=\mathbf{x}'+L$, 那么$\mathbf{x}-\mathbf{x}'=\mathbf{y}\in L$, 且$\mathcal{A}\mathbf{x}-\mathcal{A}\mathbf{x}'=\mathcal{A}(\mathbf{x}-\mathbf{x}')=\mathcal{A}(\mathbf{y})\in L$(因为$L$对$\mathcal{A}$有不变性). 从而$\mathcal{A}\mathbf{x}+L=\mathcal{A}\mathbf{x}'+L$. 如果$L$对于$\mathcal{A}$没有不变性, 那么定义商算子$\overline{\mathcal{A}}$就失去任何意义了.

设$V=L\oplus M$是个相对于$\mathcal{A}$不变的子空间的直和. 那么, 我们知道, $\mathcal{A}=\mathcal{A}_L\dot{+}\mathcal{A}_M$是$\mathcal{A}$在$L$和$M$上的限制算子的直和. 如果$f:\mathbf{u}\mapsto\mathbf{u}+L$是$M$与$\overline{V}=V/L$之间的同构(第1章§2定理10), 那么

$$(f\cdot\mathcal{A}_M)\mathbf{y}=f(\mathcal{A}_M\mathbf{y})=\mathcal{A}_M\mathbf{y}+L=\overline{\mathcal{A}}(\mathbf{y}+L)=\overline{\mathcal{A}}(f\mathbf{y}),$$

从而有

$$f\cdot\mathcal{A}_M=\overline{\mathcal{A}}\cdot f. \tag{17}$$

也就是说, $\overline{\mathcal{A}}$在$\overline{V}$上的作用与$\mathcal{A}_{\overline{M}}$在$M$上的作用是一致的, 所以就称等式(17)建立了$\overline{\mathcal{A}}$与$\mathcal{A}_M$之间的**等价性(相似)**.

例　众所周知, 作用在2维向量空间$V=\langle\mathbf{e}_1,\mathbf{e}_2\rangle$上且$\mathbf{e}_1$是个特征向量的所有的线性算子$\mathcal{B}$必对应三角形矩阵

$$\begin{pmatrix}\lambda & \sigma\\ 0 & \mu\end{pmatrix}.$$

在$\overline{V}=V/\langle\mathbf{e}_1\rangle=\langle\overline{\mathbf{e}}_2\rangle$中, 我们有$\overline{\mathcal{A}}\overline{\mathbf{e}}_2=\mu\overline{\mathbf{e}}_2$.

现在, 如果$V=\langle\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\rangle$是$\mathbb{C}$上的3维向量空间, $\mathcal{A}:V\to V$是个有特征向量$\mathbf{e}_1$($\mathcal{A}\mathbf{e}_1=\alpha\mathbf{e}_1,\alpha\in\mathbb{C}$)的线性算子. 那么, 在2维空间$\overline{V}=V/\langle\mathbf{e}_1\rangle=\langle\overline{\mathbf{e}}_2,\overline{\mathbf{e}}_3\rangle$上按前面叙述得到的商算子$\overline{\mathcal{A}}$在适当的基底之下对应三角形矩阵

$$\begin{pmatrix}\beta & \delta\\ 0 & \gamma\end{pmatrix}.$$

为简便计, 假设$(\overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3)$就是这样的基底, 于是

$$\overline{\mathcal{A}}\overline{\mathbf{e}_2} = \beta\overline{\mathbf{e}_2} \Leftrightarrow \mathcal{A}\mathbf{e}_2 = \beta\mathbf{e}_2 + \nu\mathbf{e}_1,$$

$$\overline{\mathcal{A}}\overline{\mathbf{e}_3} = \gamma\overline{\mathbf{e}_3} + \delta\overline{\mathbf{e}_2} \Leftrightarrow \mathcal{A}\mathbf{e}_3 = \gamma\mathbf{e}_3 + \delta\mathbf{e}_2 + \mu\mathbf{e}_1.$$

这样一来, 就有

$$A = \begin{pmatrix} \alpha & \nu & \mu \\ 0 & \beta & \delta \\ 0 & 0 & \gamma \end{pmatrix}.$$

可将类似的想法用到一般情形.

习 题

1. 设$\{A_i \,|1 \leqslant i \leqslant m-1\}$是个正交的等方矩阵组(见第1目结尾的说明). 证明, 对于$A = A_1 + \cdots + A_{m-1}$有$A^2 = A$, $AA_i = A_iA = A_i$, $1 \leqslant i \leqslant m-1$. 并且, 如果设$A_m = E - A$, 那么$\{A_i \,|1 \leqslant i \leqslant m\}$就是一个完全正交组.

2. 设$\mathcal{D} : M_n(\mathfrak{K}) \to M_n(\mathfrak{K})$是一个方阵空间上不恒等于$\mathcal{O}$的线性算子, 它对所有$A, B \in M_n(\mathfrak{K})$具有可乘性

$$\mathcal{D}(AB) = \mathcal{D}(A)\mathcal{D}(B).$$

证明, 必有非退化矩阵C使得$\mathcal{D} = f_c$(见§1习题3).

3. 设$\mathcal{A} : V \to V$是一个线性算子, 它对某个自然数p有$\mathrm{Im}\ \mathcal{A}^p = \mathrm{Im}\ \mathcal{A}^{p+1}$. 证明, 在这种情形$V = \mathrm{Ker}\ \mathcal{A}^p \oplus \mathrm{Im}\,\mathcal{A}^p$是$\mathcal{A}$的两个不变子空间的直和.

4. 证明, 如果作用在n维向量空间V上的线性算子$\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2, \cdots, \mathcal{A}^{n-1}$是线性无关的, 那么, 必有向量$\mathbf{v} \in V$ 使得

$$V = \langle V, \mathcal{A}\mathbf{v}, \cdots, \mathcal{A}^{n-1}\mathbf{v} \rangle$$

(此时, 说V是**循环的**).

5. 设A是$n \times n$的实矩阵但是没有实的特征值, 特别地, n为偶数且A是可逆的. 试指明, 必存在实矩阵B使得$AB = BA$, 且$B^2 = -E$, E是单位矩阵(Д. 杰考维奇(Д.Джокович))

6. 证明, 对任意A, $B \in M_n(\mathbb{R})$, 矩阵AB和矩阵BA的特征多项式相同.

7. 利用极易验证的关系式

$$A = a_0E + a_1B + a_2B^2, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

求出循环矩阵

$$A = \begin{pmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{pmatrix},$$

的特征根.

8. 证明, 半幻方空间$\mathrm{SMag}_n(\mathbb{Q})$ (见第1章§1之例8)是$M_n(\mathbb{Q})$的一个$\mathbb{Q}$-子代数.

9. [Amer. Math. Monthly. 1991年2月, p131—p133] 设A, $B \in M_n(\mathfrak{K})$, char $\mathfrak{K} \neq 2$. 若有$D = \mathrm{diag}(\theta_1, \cdots, \theta_n), \theta_i = \pm 1$, 使得$B = DA$, 则记为$B \sim A$. 显然, $\sim$是一个等价关系. 如果$S(A)$是以A为代表的等价类, 那么, Card $S(A) \leqslant 2^n$.

证明. 至少有一个$S(A)$中的矩阵, 它不以1为其特征值.

10. 设$\mathcal{A}$是个n维空间V上的线性算子, 证明,

$$\mathcal{A}^2 = \mathcal{A} \Leftrightarrow \mathrm{rank}\, \mathcal{A} + \mathrm{rank}(\mathcal{E} - \mathcal{A}) = n.$$

§4 若尔当标准型

要弄清楚一个给定的线性算子$\mathcal{A}$: $V \to V$的作用, 自然地要按照自己的目的设法在V中找一个以最好的方式适应$\mathcal{A}$的基底. 换句话来说, 需要在$\mathcal{A}$对应的相似矩阵的等价类$C^{-1}AC$中找出一个矩阵使其具有尽可能简单的形式. 根据明显的理由, 这个问题的存在与定义这个向量空间V的基础域$\mathfrak{K}$有关系. 下面, 我们认为$\mathfrak{K} = \mathbb{C}$是复数域, 尽管原则上可以用任意代数封闭域来代替$\mathbb{C}$.

1. 哈密顿–凯莱定理. 下面的并不复杂的论断是非常有用的.

定理1 线性算子$\mathcal{A}$的矩阵总能(在相似意义下)转化成三角形式.

证明 为了深信这一点, 我们用归纳法讨论. 据§3定理9, 向量空间V包含一个相对于$\mathcal{A}$的不变的超平面U : $\mathcal{A}U \subset U$. 按归纳法假设, 在U中可以选出这样的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_{n-1})$使得$\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i + \mathbf{v}_i$, $\mathbf{v}_i \in \langle \mathbf{e}_1, \cdots, \mathbf{e}_{i-1} \rangle$. 同时, $V = \langle U, \mathbf{e}_n \rangle$, 其中$\mathbf{e}_n$是任意的一个不属于$U$的向量. 设$\mathcal{A}\mathbf{e}_n = \lambda_n \mathbf{e}_n + \mathbf{u}, \mathbf{u} \in U$. 于是, 在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_{n-1}, \mathbf{e}_n)$ 之下算子$\mathcal{A}$的矩阵可以表达成所需要的形式的矩阵

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}. \square \tag{1}$$

现在可以足够简洁地证明内容丰富的

定理2(哈密顿–凯莱定理) 线性算子$\mathcal{A}$和它(在任意基底之下)对应的矩阵A必然零化自己的特征多项式, $\chi_{\mathcal{A}}(t)$, 也就是

$$\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}.$$

证明 因为这个命题与基底的选择没有关系(见§2第3目), 所以, 自然地可用定理1, 一开始, 我们就认为矩阵A在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下取如(1)的三角形式. 研究$\mathcal{A}$不变的子空间链

$$V = V_0 \supset V_1 \supset \cdots \supset V_{n-1} \supset V_n = \mathbf{0},$$

其中, $V_k = \langle \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_{n-k-1}, \mathbf{e}_{n-k} \rangle$. 因为$(\mathcal{A} - \lambda_{n-k}\mathcal{E})\, \mathbf{e}_{n-k} \in V_{k+1}$, 所以

$$(\mathcal{A} - \lambda_{n-k}\mathcal{E})\, V_k \subset V_{k+1}.$$

由此可见,

$$\begin{aligned}\chi_{\mathcal{A}}(\mathcal{A})V &= \prod_{i=1}^{n} (\mathcal{A} - \lambda_i \mathcal{E})V \\ &= (\mathcal{A} - \lambda_1\mathcal{E}) \cdots (\mathcal{A} - \lambda_n \mathcal{E})V_0 \subset (\mathcal{A} - \lambda_1\mathcal{E}) \cdots (\mathcal{A} - \lambda_{n-1}\mathcal{E})V_1 \\ &\quad \subset (\mathcal{A} - \lambda_1\mathcal{E}) \cdots (\mathcal{A} - \lambda_{n-2}\mathcal{E})V_2 \subset \cdots \subset (\mathcal{A} - \lambda_1\mathcal{E})V_{n-1} = \mathbf{0}.\end{aligned}$$

但是,$\chi_{\mathcal{A}}(\mathcal{A})V = \mathbf{0} \Leftrightarrow \chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. □

推论 线性算子$\mathcal{A}$的极小多项式$\mu_{\mathcal{A}}(t)$是其特征多项式$\chi_{\mathcal{A}}(t)$的一个因子, 而$\chi_{\mathcal{A}}(t)$可以被所有的线性因子$t - \lambda$整除, $\lambda \in \mathrm{Spec}(\mathcal{A})$.

证明 按定义$\mu_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$, 据定理2, 又有$\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. 由§2之定理2推出$\mu_{\mathcal{A}}(t)$整除$\chi_{\mathcal{A}}(t)$.

其次, 如果λ是算子$\mathcal{A}$的一个特征值, 那么,

$$\mathcal{A}\mathbf{v} = \lambda \mathbf{v} \Rightarrow \mathbf{0} = \mu_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mu_{\mathcal{A}}(\lambda)\mathbf{v} \Rightarrow \mu_{\mathcal{A}}(\lambda) = 0 \Rightarrow (t - \lambda) | \mu_{\mathcal{A}}(t)$$

(我们反复使用了§3中蕴涵式(7)的结论). □

说明 似乎, 哈密顿–凯莱定理就是证明了$\det(tE - A)|_{t=A} = \det(AE - A) = \det\ 0 = 0$.但这是个完全不可信的推理. 想一想, 为什么.

哈密顿–凯莱定理有很多应用, 但我们暂且只把它应用到极直接的形式中.

例1 设$\mathcal{A}$: $V \to V$是个幂零指数为m的幂零的线性算子(见§2的定义3), 所以, $\mu_{\mathcal{A}}(t) = t^m$.

设$\mathcal{A}^{m-1}\mathbf{v} \neq \mathbf{0}$. 那么, 向量$\mathbf{v}$, $\mathcal{A}\mathbf{v}, \cdots, \mathcal{A}^{m-1}\mathbf{v}$是线性无关的.

事实上, 所有的非平凡的线性关系式必形如

$$\mathcal{A}^k\mathbf{v} + \alpha_1 \mathcal{A}^{k+1}\mathbf{v} + \cdots + \alpha_{m-1-k}\mathcal{A}^{m-1}\mathbf{v} = \mathbf{0}, \quad 0 \leqslant k \leqslant m-1.$$

把算子$\mathcal{A}^{m-1-k}$作用到这个等式的两侧, 即可确信$\mathcal{A}^{m-1}\mathbf{v} = \mathbf{0}$; 这与$V$的选择相矛盾.

也就是说, 由哈密顿–凯莱定理推出, 自然地, 算子$\mathcal{A}$的幂零指数m不超过$n = \dim\ V$. 现在, 设$m = n$且$\mathcal{A}^{n-1}\mathbf{e} \neq \mathbf{0}$, 就可导出下面对基底向量的约定

$$\mathbf{e}_1 = \mathcal{A}^{n-1}\mathbf{e}, \quad \mathbf{e}_2 = \mathcal{A}^{n-2}\mathbf{e}, \cdots, \mathbf{e}_{n-1} = \mathcal{A}\mathbf{e}, \quad \mathbf{e}_n = \mathbf{e}.$$

于是$\mathcal{A}\mathbf{e}_1 = \mathbf{0}$, $\mathcal{A}\mathbf{e}_k = \mathbf{e}_{k-1}, k > 1$, 而且算子$\mathcal{A}$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的矩阵就是下面将要给出定义的$\lambda$=0的若尔当块$J_n(\lambda)$.

如果, 再设$V = \langle 1, t, \cdots, t^{n-1} \rangle$是$\mathbb{C}$上次数$<n$的多项式空间而$\mathcal{D} = \dfrac{d}{dt}$是对$t$求导数的算子, 那么, 这个算子在基底$(\mathbf{e}_i)$, $\mathbf{e}_i = \dfrac{1}{i!}t^i$之下的矩阵刚好就是$J_n(0)$.

定义1 i) 我们称[1]矩阵

$$J_m(\lambda)=\begin{pmatrix}\lambda & 1 & 0 & \cdots & 0 & 0\\ 0 & \lambda & 1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda & 1\\ 0 & 0 & 0 & \cdots & 0 & \lambda\end{pmatrix}$$

为特征值λ对应的$m\times m$阶(或m阶)的**(上)若尔当块**.

ii) 一个矩阵, 若它的对角线由若尔当块组成, 且这些块之外均为0:

$$J=\begin{pmatrix}J_{m_1}(\lambda_1) & \cdots & 0\\ \cdots & \cdots & \cdots\\ 0 & \cdots & J_{m_s}(\lambda_s)\end{pmatrix}, \tag{2}$$

则被称为**若尔当矩阵**.

iii) 线性算子$\mathcal{A}$: $V\to V$, 在V的一个基底之下, 对应的矩阵A是若尔当矩阵, 或者如常说的, 有**若尔当标准型**$J(\mathcal{A})$, 则称这个基底为**若尔当基底**.

vi) 把求解矩阵方程式$X^{-1}AX=J(A)$称之为把矩阵A化成**若尔当标准型**, 其中X是(未知的)非退化矩阵, 而$J(A)$是(未知的)若尔当矩阵.

说明 $J_m(\lambda)-\lambda E=J_m(0)$是个幂零矩阵. 特别地, $(t-\lambda)^m$是若尔当矩阵(2)的极小多项式, 且λ是它的唯一的特征值: Spec $J_m(\lambda)=\{\lambda\}$.

例2 设$D_n(\lambda)$是所有形如$e^{\lambda t}f(t)$的复函数的向量空间, 其中$\lambda\in\mathbb{C}$, $f(t)$取遍所有次数$\leqslant n-1$的多项式. 因为,

$$\frac{d}{dt}(e^{\lambda t}f(t))=e^{\lambda t}(\lambda f(t)+f'(t)),$$

所以, 求导数$\mathcal{D}=\dfrac{d}{dt}$是$D_n(\lambda)$上的一个线性算子. 设$\mathbf{e}_{i+1}=\dfrac{t^i}{i!}e^{\lambda t}, i=0,\cdots,n-1$. 显然,

$$\mathcal{D}\mathbf{e}_{i+1}=\frac{t^{i-1}}{(i-1)!}e^{\lambda t}+\lambda\frac{t^i}{i!}e^{\lambda t}=\mathbf{e}_i+\lambda\mathbf{e}_{i+1}$$

(0!=1; 当$i=0$时, 第一个被加项就没有了). 于是; 函数$\dfrac{t^i}{i!}e^{\lambda t}$在我们的向量空间中对于线性算子$\mathcal{D}$构成一个若尔当基底, 且$J(\mathcal{D})=J_n(\lambda)$.

这个例子指明了若尔当矩阵在线性微分方程理论中的特殊作用, 我们再次回到它上面.

---

1) 为纪念法国数学家C. Jordan(1838—1922)

例3 如果$f(t)$是任意一个多项式, 那么

$$f(J_m(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda)/1! & f''(\lambda)/2! & \cdots & f^{(m-1)}(\lambda)/(m-1)! \\ 0 & f(\lambda) & f'(\lambda)/1! & \cdots & f^{(m-2)}(\lambda)/(m-2)! \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & f(\lambda) \end{pmatrix},$$

因此, 运算$J_m(\lambda)$比运算一般矩阵要容易得到.

2. 若尔当标准型: 定理与推论 我们阐述基本命题及其结论.

基本定理 *代数闭域上(特别地, $\mathbb{C}$上)的任意一个n阶方阵A都可以化成若尔当标准型. 即存在非退化矩阵C使得$C^{-1}AC = J(A) = J$是个形如(2)的矩阵. 不计小块之间的置换, 矩阵的若尔当标准型是唯一的.*

因为相似的矩阵的极小多项式是相同的, 所以由基本定理及关于若尔当块$J_m(\lambda)$所做的说明, 得到

$$\mu_A(t) = (t-\lambda_{i_1})^{m_{i_1}} \cdots (t-\lambda_{i_p})^{m_{i_p}}, \qquad (3)$$

其中$\{\lambda_{i_1}, \cdots, \lambda_{i_p}\}$是$A$的所有的两两不同的特征值, 而$m_{j_k}$是相对于特征值$\lambda_{j_k}$的若尔当块的最大阶数.

显然, 矩阵A可对角化(也就是它相似于矩阵$\mathrm{diag}(\lambda_1, \cdots, \lambda_n)$)的充要条件是在$J(A)$ 中不存在>1阶的若尔当块. 进而再考虑到(3)就得到下面的一个有用的判别法.

推论 *域$\mathbb{C}$上方阵A可化成对角型, 当且仅当, 它的极小多项式$\mu_A(t)$没有重根.*

这个推论是有效用的, 因为要计算极小多项式$\mu_A(t)$并不一定要把A化成若尔当标准型.

基本定理的证明将分成三个部分, 分别在第3至第5目. 还要顺便讲述一些求若尔当标准型的可实践的建议, 其后, 再指明一个另外的证明.

3. 根子空间 我们引进下面的

定义2 *向量集合*

$$V(\lambda) = \{\mathbf{v} \in V \mid \text{对某个}k, (\mathcal{A}-\lambda\mathcal{E})^k\mathbf{v} = \mathbf{0}\}$$

*被称为特征值$\lambda \in \mathrm{Spec}(\mathcal{A})$对应的***根子空间**.

关于$V(\lambda)$是个子空间, 我们很容易相信这是正确的. 例如, 取$\mathbf{u} \in V(\lambda), \mathbf{v} \in V(\lambda)$,其中$(\mathcal{A}-\lambda\mathcal{E})^s\mathbf{u} = \mathbf{0}$, $(\mathcal{A}-\lambda\mathcal{E})^t\mathbf{v} = \mathbf{0}$; 且$m = \max\{s, t\}$, 那么

$$(\mathcal{A}-\lambda\mathcal{E})^m(\alpha\mathbf{u}+\beta\mathbf{v}) = \alpha(\mathcal{A}-\lambda\mathcal{E})^m\mathbf{u} + \beta(\mathcal{A}-\lambda\mathcal{E})^m\mathbf{v} = \mathbf{0},$$

从而, 对任意$\alpha, \beta \in \mathbb{C}$有$\alpha\mathbf{u}+\beta\mathbf{v} \in V(\lambda)$. 因为, $V(\lambda)$包含λ对应的特征向量, 所以$V(\lambda) \neq \mathbf{0}$. 其次, $V^\lambda \subset V(\lambda)$, 正如幂零指数$n > 1$的算子$\mathcal{A}$的例子表明的那样,

不等号是可能成立的. 在这种情形, $\lambda=0$是它的唯一的一个特征值, $\dim V^0 = 1$, 但$V(0)=V$.

因为$\dim(\lambda) \leqslant n$, 且$\mathcal{A}-\lambda\mathcal{E}$在$V(\lambda)$上的限制是个幂零算子, 所以

$$V(\lambda) = \{\mathbf{v} \in V | (\mathcal{A}-\lambda\mathcal{E})^n = \mathbf{0}\}.$$

定理3　设$\mathcal{A}: V \to V$是以

$$\chi_{\mathcal{A}}(t) = \prod_{i=1}^{p}(t-\lambda_i)^{n_i}; \quad \lambda_i \neq \lambda_j \ npu\ i \neq j.$$

为特征多项式的线性算子. 那么, $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$是根子空间$V(\lambda_i)$的直和, 它们之间的每一个都是对$\mathcal{A}$不变的, 而且具有维数$\dim V(\lambda_i) = n_i$. 算子$\mathcal{A}-\lambda_i\mathcal{E}$在$V(\lambda_i)$上是幂零的, 且以非退化方式作用在子空间

$$V_i = V(\lambda_1) \oplus \cdots \oplus V(\lambda_{i-1}) \oplus V(\lambda_{i+1}) \oplus \cdots \oplus V(\lambda_p).$$

上. 最后, λ_i是算子$\mathcal{A}|_{V(\lambda_i)}$的唯一的一个特征值.

证明　没有一个素因式$t-\lambda_k$能够同时是所有的多项式

$$\chi_i(t) = \prod_{j\neq i}(t-\lambda_j)^{n_j}, \quad i=1,2,\cdots,p$$

的因子, 所以$(\chi_1(t),\cdots,\chi_p(t))$的最高公因式是1. 从而可找到多项式$f_1(t),\cdots,f_p(t) \in \mathbb{C}[t]$, 使得

$$\sum_{i=1}^{p}\chi_i(t)f_i(t) = 1. \tag{4}$$

子空间

$$W_i = \chi_i(\mathcal{A})f_i(\mathcal{A})V = \{\chi_i(\mathcal{A})f_i(\mathcal{A})\mathbf{v} | \mathbf{v} \in V\}, \quad 1 \leqslant i \leqslant p,$$

对于$\mathcal{A}$是不变的:

$$\mathcal{A}W_i = \chi_i(\mathcal{A})f_i(\mathcal{A})\mathcal{A}V \subset \chi_i(\mathcal{A})f_i(\mathcal{A})V = W_i.$$

此外.

$$(\mathcal{A}-\lambda_i\mathcal{E})^{n_i}W_i = \chi_{\mathcal{A}}(\mathcal{A})f_i(\mathcal{A})V = \mathbf{0}$$

(因为, 据定理2, $\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$), 所以,

$$W_i \subset V(\lambda_i). \tag{5}$$

可以写成形如

$$\mathcal{E} = \sum_{i=1}^{p}\chi_i(\mathcal{A})f_i(\mathcal{A}),$$

的关系(4)给我们提供了一个分解

$$V=\sum_{i=1}^{p}W_i$$

以及更一般的(借助包含式(5))

$$V=\sum_{i=1}^{p}V(\lambda_i).$$

我们假设$\mathbf{v}\in V(\lambda_i)\cap V_i$, 其中, 如定理叙述的那样, $V_i=\sum_{j\neq i}V(\lambda_j)$. 从而$(\mathcal{A}-\lambda_i\mathcal{E})^n\mathbf{v}=\mathbf{0}$, 又因为$\mathbf{v}=\sum_{j\neq i}\mathbf{v}_j$而且$(\mathcal{A}-\lambda_j\mathcal{E})^n\mathbf{v}_j=\mathbf{0}$, 所以有$\left\{\prod_{j\neq i}(\mathcal{A}-\lambda_j\mathcal{E})^n\right\}\mathbf{v}=\mathbf{0}$. 但是, 由多项式$(t-\lambda_i)^n$与$c(t)=\prod_{j\neq i}(t-\lambda_j)^n$的互素性可以推出, 存在$a(t),b(t)$ 使得

$$a(t)(t-\lambda_i)^n+b(t)c(t)=1.$$

我们就得到

$$\mathbf{v}=a(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{E})^n\mathbf{v}+b(\mathcal{A})\left\{\prod_{j\neq i}(\mathcal{A}-\lambda_j\mathcal{E})^n\right\}\mathbf{v}=\mathbf{0},$$

也就是说向量空间$V(\lambda_i)$和V_i不相交. 这就意味着, 有$\mathcal{A}$的不变子空间的直和

$$V=V(\lambda_1)\oplus\cdots\oplus V(\lambda_p)\tag{6}$$

分解.

由包含式(5)和分解式(6)可直接到$W_i=V(\lambda_i)$. 这样一来, 对于$V(\lambda_i)$就得到一个有效的表达式

$$V(\lambda_i)=\chi_i(\mathcal{A})f_i(\mathcal{A})V,$$

其中$\chi_i(t)$, $f_i(t)$是恒等(4)中的多项式. 特别地,

$$(\mathcal{A}-\lambda_i\mathcal{E})^nV(\lambda_i)=\mathbf{0}.$$

$\mathcal{A}$在$V(\lambda_i)$上的极小多项式是多项式$(t-\lambda_i)^{n_i}$的某一个因式. 由此得到, 首先, λ_i是算子$\mathcal{A}|_{V(\lambda_i)}$的唯一一个特征值. 其次, 在子空间$V(\lambda_i)$的基底合并成的基底之下, 算子$\mathcal{A}$有矩阵

$$A=\begin{pmatrix}A_1\cdots & 0\\ \cdots\cdots\cdots & \\ 0\ \cdots & A_p\end{pmatrix},$$

其中A_i是唯一特征值λ_i对应的$n'_i = \dim\ V(\lambda_i)$阶的矩阵, 而且, 它的特征多项式是

$$\chi_{A_i}(t) = (t-\lambda_i)^{n'_i}, \quad n'_i \leqslant n_i.$$

因为$\chi_A(t) = \prod_{i=1}^{p} \chi_{A_i}(t)$, 所以, $n = n'_1 + \cdots + n'_p$且$n'_i = n_i$.

剩下来是要证明限制算子$(\mathcal{A}-\lambda_i\mathcal{E})|_{V_i}$的非退化性. 但, 这是显然的, 在相反的情形$\{\mathrm{Ker}(\mathcal{A}-\lambda_i\mathcal{E})\} \cap V_i \neq \mathbf{0}$, 而且必有$\mathbf{0} \neq \mathbf{v} \in V_i$使得$\mathcal{A}\mathbf{v} - \lambda_i\mathbf{v} = \mathbf{0}$.但是在$V_i$上$\mathcal{A}$的特征多项式是$\chi_i(t) = \prod_{j\neq i}(t-\lambda_j)^{n_j}$, 而$\lambda_i$绝不可能是它的特征值. □

4. 幂零算子的情形 定理3把选择线性算子$\mathcal{A}$: $U \to V$的最简矩阵的问题归结为那样的情形, 即$\mathcal{A}$有唯一的特征值λ且$(\mathcal{A}-\lambda\mathcal{E})^m = \mathcal{O}$的情形, $m \leqslant \dim\ V$. 设$\mathcal{B} = \mathcal{A} - \lambda\mathcal{E}$, 我们就得到一个幂零指数为$m$的幂零算子, 它对应幂零矩阵$B$. 要研究的情况是很自然的

定义3 称线性包络

$$\mathfrak{K}[\mathcal{B}]\,\mathbf{v} = \langle \mathbf{v}, \mathcal{B}\mathbf{v}, \cdots, \mathcal{B}^{m'-1}\mathbf{v}\rangle$$

是与幂零指数为m的算子$\mathcal{B}$和向量$\mathbf{v}$相关联的**循环子空间**. 设$m' \leqslant m$是使$\mathcal{B}^{m'}\mathbf{v} = \mathbf{0}$的最小的自然数.

定理4 幂零矩阵B的若尔当标准型$J(B)$存在(基础域$\mathfrak{K}$是任意的).

证明 由例1和定义1可以看出, 所有的循环子空间都对应若尔当块. 我们需要指出, 幂零算子$\mathcal{B}$和幂零矩阵B所作用的向量空间V可以分解成用适当方法选出的循环子空间的直和.

据定理1, 矩阵B可以转化为对角线上元素为0的上三角形矩阵. 这意味着, 前$n-1$个基底向量包络U对$\mathcal{B}$是不变的. 按定义$\mathcal{B}V \subset U$, 而按归纳假定, 在U中可以选择出相对于$\mathcal{B}$的若尔当基底, 或者, 同样的

$$\begin{gathered} U = \mathfrak{K}[\mathcal{B}]\,\mathbf{e}_1 \oplus \cdots \oplus \mathfrak{K}[\mathcal{B}]\,\mathbf{e}_s, \\ \mathfrak{K}[\mathcal{B}]\,\mathbf{e}_i = \left\langle \mathbf{e}_i, \mathcal{B}\mathbf{e}_i, \cdots, \mathcal{B}^{m_i-1}\mathbf{e}_i \right\rangle, \quad \mathcal{B}^{m_i}\mathbf{e}_i = \mathbf{0}. \end{gathered} \tag{7}$$

不失一般性, 可以认为

$$m_1 \geqslant m_2 \geqslant \cdots \geqslant m_s. \tag{8}$$

其次, 对任意一个不含在U中的向量$\mathbf{v}$, 有$V = \langle \mathbf{v}, U\rangle$, $\mathcal{B}\mathbf{v} \in U$, 所以, $\mathcal{B}\mathbf{v} = \sum_i \alpha_i\mathbf{e}_i + \mathcal{B}\mathbf{u}$, $\mathbf{u} \in U$. 用$\mathbf{v}' = \mathbf{v} - \mathbf{u}$替换$\mathbf{v}$, 我们得到

$$V = \langle \mathbf{v}', U\rangle, \quad \mathcal{B}\mathbf{v}' = \sum_{i=1}^{s} \alpha_i\mathbf{e}_i.$$

如果$\alpha_i = 0$, $1 \leqslant i \leqslant s$, 那么, 把对应的循环子空间$\langle \mathbf{v}'\rangle$的$J_1(0)$补充到若尔当块$J_{m_1}(0), \cdots, J_{m_s}(0)$, 也就是

$$B \sim J(B) = \mathrm{diag}(J_{m_1}(0), \cdots, J_{m_s}(0), J_1(0))$$

(~是相似号).

剩下来要研究对某个指标$r \geqslant 1$,

$$\alpha_1 = \cdots = \alpha_{r-1} = 0, \quad \mathcal{B}\mathbf{v}' = \sum_{i=r}^{s} \alpha_i \mathbf{e}_i, \quad \alpha_r \neq 0$$

的情形. 为方便计, 设

$$\mathbf{e}_i' = \mathbf{e}_i, \quad i \neq r, \quad \mathbf{e}_r' = \frac{1}{\alpha_r}\mathbf{v}', \quad \beta_i = \frac{\alpha_i}{\alpha_r}.$$

于是,

$$\mathcal{B}\mathbf{e}_r' = \mathbf{e}_r + \sum_{i=r+1}^{s} \beta_i \mathbf{e}_i := \mathbf{f}_r.$$

与顺序关系式(8)相对应, $\mathcal{B}^{m_r}\mathbf{f}_r = \mathbf{0}$. 又因为(7)是直和, 所以, 无论怎样系数$\beta_i$都必有$\mathcal{B}^{m_r-1}\mathbf{f}_r \neq \mathbf{0}$. 此外, 一个很容易的推断就可以表明, 和

$$\sum_{i \neq r} \mathfrak{K}[\mathcal{B}]\,\mathbf{e}_i' + \mathfrak{K}[B]\,\mathbf{f}_r$$

同样是直和并且与U重合.

但现在, 循环子空间$\mathfrak{K}[\mathcal{B}]\,\mathbf{f}_r$被$\mathbf{e}_r' \notin U$扩大了: $\mathfrak{K}[\mathcal{B}]\,\mathbf{f}_r \subset \mathfrak{K}[\mathcal{B}]\,\mathbf{e}_r'$, 从而我们有直和

$$V = \bigoplus_{i=1}^{s} \mathfrak{K}[\mathcal{B}]\,\mathbf{e}_i',$$

对应一组指数$m_1', \cdots, m_s'$, 其中$m_i' = m_i$, $i \neq r$, $m_r' = m_r + 1$. 同样也有

$$B \sim \mathrm{diag}(J_{m_1'}(0), \cdots, J_{m_s'}(0))$$

(若尔当块的个数保持了原来的个数, 但有一小块的维数增加了个1). 一般说来, 序列$(m_1', \cdots, m_s')$是没有次序的, 但通过向量$\mathbf{e}_i'$的变换就可以达到要求. 这样一来, 对于幂零算子$\mathcal{B}$的若尔当基底就证明了它的存在性. □

5. 唯一性 要着手证明唯一性, 我们顺便讲一个把n阶方阵化成若尔当标准型的有实践意义的原则.

为此, 需要会找出与矩阵A的特征值λ相对应的m阶若尔当块$J_m(\lambda)$的个数$N(m, \lambda)$. 用通常方式, 把作用在n维向量空间V上的线性算子$\mathcal{A}$的矩阵A与V的直和分解

$$V = V(\lambda) \oplus V', \tag{9}$$

一起讨论. 其中

$$V(\lambda) = \bigoplus_{j=1}^{s} \left\langle \mathbf{e}_j, (\mathcal{A} - \lambda\mathcal{E})\mathbf{e}_j, \cdots, (\mathcal{A} - \lambda\mathcal{E})^{m_j-1}\mathbf{e}_j \right\rangle,$$

$$V' = \sum_{\lambda' \neq \lambda} V(\lambda').$$

我们先来计算矩阵$(A-\lambda E)^t$ 的秩$r_t=\operatorname{rank}(A-\lambda E)^t$, 或者, 同样的, 空间$(\mathcal{A}-\lambda\mathcal{E})^tV$的维数. 实际上, 这个维数与$V$的基底的选择没有关系. 在分解式(9)中, 每个空间对$(\mathcal{A}-\lambda\mathcal{E})^t$ 都是不变的, 所以

$$\dim(\mathcal{A}-\lambda\mathcal{E})^tV=\sum_j\dim(\mathcal{A}-\lambda\mathcal{E})^t\mathbb{C}[\mathcal{A}]\,\mathbf{e}_j+\dim(\mathcal{A}-\lambda\mathcal{E})^tV'.$$

为确定起见, 可以认为$m_1\leqslant m_2\leqslant\cdots\leqslant m_s$. 如果$m_j\leqslant t$, 那么, $(\mathcal{A}-\lambda\mathcal{E})^t\mathbb{C}[\mathcal{A}]\,\mathbf{e}_j=\mathbf{0}$. 当$m_j>t$ 时, 我们有

$$(\mathcal{A}-\lambda\mathcal{E})^t\mathbb{C}[\mathcal{A}]\,\mathbf{e}_j=\left\langle(\mathcal{A}-\lambda\mathcal{E})^t\mathbf{e}_j,(\mathcal{A}-\lambda\mathcal{E})^{t+1}\mathbf{e}_j,\cdots,(\mathcal{A}-\lambda\mathcal{E})^{m_j-1}\mathbf{e}_j\right\rangle,$$

也就是

$$\dim(\mathcal{A}-\lambda\mathcal{E})^t\mathbb{C}[\mathcal{A}]\,\mathbf{e}_j=m_j-t.$$

在V' 上算子$(\mathcal{A}-\lambda\mathcal{E})$ 是非退化的(定理1), 所以,

$$\dim(\mathcal{A}-\lambda\mathcal{E})^tV'=\dim V'.$$

我们就得到

$$r_t=\sum_{m_j>t}(m_j-t)+\dim V',$$

从而

$$\begin{aligned}r_t-r_{t+1}&=\sum_{m_j>t}(m_j-t)-\sum_{m_j>t+1}(m_j-t-1)\\&=\sum_{m_j>t}(m_j-t)-\sum_{m_j>t+1}(m_j-t)+\sum_{m_j>t+1}1\\&=\sum_{m_j=t+1}1+\sum_{m_j>t+1}1=N(t+1,\lambda)+N(t+2,\lambda)+\cdots,\end{aligned}$$

也就是

$$\begin{aligned}&r_{m-1}-r_m-(r_m-r_{m+1})\\&=\{N(m,\lambda)+N(m+1,\lambda)+\cdots\}-\{N(m+1,\lambda)+N(m+2,\lambda)+\cdots\}\\&=N(m,\lambda),\end{aligned}$$

进而, 我们得到了最终的公式

$$N(m,\lambda)=r_{m-1}-2r_m+r_{m+1},\tag{10}$$

$$m\geqslant1,\qquad r_t=\operatorname{rank}(\mathcal{A}-\lambda\mathcal{E})^t,\quad r_0=n.$$

我们要注意, r_t 是矩阵A 的不变量(也就是, 由矩阵A 的相似类确定的数). 这就意味着, 公式(10)也就建立了若尔当型矩阵$J(A)$ 的唯一性. □

迄今为止, 关于可实现相似

$$J(A) = C^{-1}AC$$

的矩阵C 差不多是什么也没讲. 但是, 现在, 因为A 和$J(A)$ 对我们来说都是已知矩阵了, 就可以利用矩阵方程

$$XJ(A) - AX = 0$$

求出$C = (c_{ij})$, 关于这个方程可回想定义1中的iii)以及和它等价的n^2阶齐次线性方程组. 设$C_1, \cdots, C_r$ 是它的一个基础解系. 一般说来, 并非所有C_i 都是非退化矩阵, 但是, 因为若尔当标准型$J(A)$ 确实存在, 所以(带有不定系数$t_1, \cdots, t_r$ 的)

$$\det(t_1C_1 + \cdots + t_rC_r) \neq 0,$$

因此就有复数$\alpha_1, \cdots, \alpha_r$ 使得

$$\det(\alpha_1C_1 + \cdots + \alpha_rC_r) \neq 0.$$

从而$C = \alpha_1C_1 + \cdots + \alpha_rC_r$ 就是未知的矩阵. 当然, 远非这么一种方式可以确定C , 甚至可以规范$\det\ C = 1$. 虽然没有原则性困难, 但是, 用这种方式找出可以实现向若尔当基底转化的矩阵C , 并不太容易实现.

例4 满足关系式$S^2 = nS$ 的$n \times n$ 矩阵

$$S = \sum_{i,j} E_{ij} = \begin{pmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{pmatrix},$$

的极小多项式, 显然是$\mu_S(t) = t^2 - nt$, 即$\lambda_1 = n, \lambda_2 = 0$ 分别是S 的1重根和$n-1$ 重根. 其次, 因为rank $S = 1$ 且S 不可能是幂零矩阵(极小多项式不幂零), 所以它的若尔当标准型就剩下一种可能: $J(S) = \mathrm{diag}(n, 0, \cdots, 0)$. 这同样是显然的, 因为$P = \dfrac{1}{n}S$ 是个投影矩阵. 特别地,

$$C = \begin{pmatrix} 1/n & -1 & \cdots & -1 \\ 1/n & n-1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots \\ 1/n & -1 & \cdots & n-1 \end{pmatrix}.$$

就是矩阵方程

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix} \cdot \begin{pmatrix} n & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}$$

的一个解.

6. 化若尔当标准型的其他方法 为了求出矩阵的若尔当标准型和相应的若尔当基底, 主理想环上的一般模的理论相当有用, 已被深入研究的λ-矩阵方法是上述理论的一个变形. 尽管这个理论已经普及并且导出了很多其他的重要结论, 但它的内容是足够令人厌倦的, 而且这本教科书也未必能容纳得下. 相反, 我们紧接着要引入的直接的几何方法足够直观且能够推出一个联合定理3和定理4的不变量(例如, 见参考书 [2, 9]).

于是, 设$\mathcal{A}: V \to V$ 是个复的线性算子, λ是它的一个特征值. 因为对$\mathcal{A} - \lambda\mathcal{E}$ 的若尔当基底就是对于$\mathcal{A}$ 的若尔当基底, 所以, 不失一般性, 可以认为, $\lambda = 0$. 此时, 算子$\mathcal{A}$ 就是退化的. 包含式$\mathrm{Im}\ \mathcal{A} \subset V$ 是严格成立的. 因而, 我们可以对$n = \dim\ V$ 用归纳法. 为此目的, 我们看, 稳定在某个p之第p项上的序列

$$\mathrm{Im}\,\mathcal{A}^0 \supset \mathrm{Im}\,\mathcal{A}^1 \supset \mathrm{Im}\,\mathcal{A}^2 \supset \cdots \supset \mathrm{Im}\,\mathcal{A}^{p-1} \supset \mathrm{Im}\,\mathcal{A}^p = \mathrm{Im}\,\mathcal{A}^{p+1} = \cdots,$$

就有

$$\mathrm{Im}\,\mathcal{A}^p \cap \mathrm{Ker}\ \mathcal{A} = \mathbf{0}, \quad \mathrm{Im}\,\mathcal{A}^{p-1} \cap \mathrm{Ker}\mathcal{A} \neq \mathbf{0}.$$

对照§3的习题3, V 有对$\mathcal{A}$ 不变的子空间的直和分解

$$V = \mathrm{Ker}\ \mathcal{A}^p \oplus \mathrm{Im}\ \mathcal{A}^p.$$

如果$\dim\ \mathrm{Im}\ \mathcal{A}^p > 0$, 那么, 按归纳法假定, 其中的每个被加项都有若尔当基底, 把它们合并起来就是V 的一个若尔当基底.

我们进入了

$$V = \mathrm{Ker}\ \mathcal{A}^p,$$

也就是, $\mathcal{A}^p = \mathcal{O}$, 而$\mathcal{A}^{p-1} \neq \mathcal{O}$ 的情形.

这样一来, 绕开具有独立价值的定理3, 得到了简化成幂零算子的情形. 不求助于定理4, 我们以下面的方式进入. 为了方便, 令

$$V_i = \mathrm{Im}\,\mathcal{A}^{i-1} \cap \mathrm{Ker}\ \mathcal{A}.$$

那么,

$$\mathrm{Ker}\ \mathcal{A} = V_1 \supset V_2 \supset \cdots \supset V_p \neq \mathbf{0}, \quad V_{p+1} = \mathbf{0}.$$

在子空间V_p 中选择基底$(\mathbf{a}_i^1; 1 \leqslant i \leqslant n_p)$. 因为$\mathbf{a}_i^1 \in \mathrm{Im}\ \mathcal{A}^{p-1}$, 所以$\mathbf{a}_i^1 = \mathcal{A}^{p-1}\mathbf{a}_i^p$, 对某个向量$\mathbf{a}_i^p$ 成立. 我们看向量$\mathbf{a}_i^k = \mathcal{A}^{p-k}\mathbf{a}_i^p$, $1 \leqslant k \leqslant p$. 把向量组$\mathbf{a}_i^1$ 补充上向量组$\mathbf{b}_j^1$ 后扩充成V_{p-1} 的基底. 再找出$\mathbf{b}_j^{p-1}$ 使得$\mathbf{b}_j^1 = \mathcal{A}^{p-2}\mathbf{b}_j^{p-1}$, 看$\mathbf{b}_j^l = \mathcal{A}^{p-l-1}\mathbf{b}_j^{p-1}$, $1 \leqslant l \leqslant p-1$. 然后, 把$\mathbf{a}_i^1$, $\mathbf{b}_j^1$ 加上向量$\mathbf{c}_k^1$ 共计n_{p-2} 个向量一起扩充组成子空间V_{p-2}

的一个基底, 依此类推. 用一个图来说明这个过程

$$\begin{array}{llllll}
\mathbf{a}_i^p & & & & & \\
\downarrow & & & & & \\
\mathbf{a}_i^{p-1} & \mathbf{b}_j^{p-1} & & & & \\
\downarrow & \downarrow & & & & \\
\mathbf{a}_i^{p-2} & \mathbf{b}_j^{p-2} & \mathbf{c}_k^{p-2} & & & \\
\downarrow & \downarrow & \downarrow & & & \\
\cdots & \cdots & \cdots & \cdots & & \\
\downarrow & \downarrow & \downarrow & & & \\
\mathbf{a}_i^2 & \mathbf{b}_j^2 & \mathbf{c}_k^2 & \cdots & \mathbf{d}_s^2 & \\
\downarrow & \downarrow & \downarrow & \cdots & \downarrow & \\
\mathbf{a}_i^1 & \mathbf{b}_j^1 & \mathbf{c}_k^1 & \cdots & \mathbf{d}_s^1 & \mathbf{e}_t^1 \\
\downarrow & \downarrow & \downarrow & \cdots & \downarrow & \downarrow \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}.
\end{array}$$

我们假设图中的向量是线性相关的:

$$\sum_i \alpha_i^{(p)} \mathbf{a}_i^p + \sum_i \alpha_i^{(p-1)} \mathbf{a}_i^{p-1} + \cdots + \sum_j \beta_j^{(p-1)} \mathbf{b}_j^{p-1} + \cdots + \sum_t \delta_t \mathbf{e}_t = \mathbf{0}. \qquad (11)$$

把算子$\mathcal{A}^{p-1}$ 作用到(11), 即得结果:

$$\sum_i \alpha_i^{(p)} \mathbf{a}_i^1 = 0,$$

从而, 我们有

$$\alpha_i^{(p)} = 0, \quad 1 \leqslant i \leqslant n_p.$$

现在, 再把算子$\mathcal{A}^{p-2}$ 用于(11), 我们得到

$$\sum_i \alpha_i^{(p-1)} \mathbf{a}_i^1 + \sum_j \beta_j^{(p-1)} \mathbf{b}_j^1 = \mathbf{0},$$

按照向量的选择, 可得出$\alpha_i^{(p-1)} = 0 = \beta_j^{(p-1)}$. 继续这个过程就可推出(11)是个平凡关系式. 另一方面, 图中向量的总个数等于

$$\begin{aligned}
& |\{\mathbf{a}_i^k\}| + |\{\mathbf{b}_j^l\}| + \cdots + |\{\mathbf{e}_t\}| \\
= & \, p n_p + (p-1) n_{p-1} + \cdots + 2 n_2 + n_1 \\
= & \, (n_1 + \cdots + n_p) + (n_2 + \cdots + n_p) + \cdots + (n_{p-1} + n_p) + n_p \\
= & \dim V_1 + \dim V_2 + \cdots + \dim V_{p-1} + \dim V_p \\
= & \sum_{i=1}^p (\mathrm{Im}\, \mathcal{A}^{i-1} \cap \mathrm{Ker}\, \mathcal{A}) = \sum_i (\dim \mathrm{Ker}\, \mathcal{A}^i - \dim \mathrm{Ker}\, \mathcal{A}^{i-1}) \\
= & \dim \mathrm{Ker}\, \mathcal{A}^p = \dim V
\end{aligned}$$

(关于这个等式, 参见§2的习题7).

这样一来, 图中的向量就组成了空间V 的一个基底, 而且按其构作方法, 它是个若尔当基底.

谈到关于若尔当标准型唯一性的论断, 那么, 在第5目的表达式中, 我们有

$$\begin{aligned}
N(m,0) &= \dim\ V_m - \dim\ V_{m+1} \\
&= (\dim\ \mathrm{Ker}\ \mathcal{A}^m - \dim\ \mathrm{Ker}\ \mathcal{A}^{m-1}) - (\dim\ \mathrm{Ker}\ \mathcal{A}^{m+1} - \dim\ \mathrm{Ker}\ \mathcal{A}^m) \\
&= 2\dim\ \mathrm{Ker}\ \mathcal{A}^m - \dim\ \mathrm{Ker}\ \mathcal{A}^{m-1} - \dim\ \mathrm{Ker}\ \mathcal{A}^{m+1} \\
&= \mathrm{rank}\ \mathcal{A}^{m-1} - 2\ \mathrm{rank}\ \mathcal{A}^m + \mathrm{rank}\ \mathcal{A}^{m+1} = r_{m-1} - 2r_m + r_{m+1},
\end{aligned}$$

而这个数, 正如我们知道的, 它是个不变量.

7. 其他的标准型 在这一目, 我们简要地叙述矩阵的一些另外的标准型, 特别地, 它们适用于不是代数封闭域的情形.

i) 循环矩阵和循环块. 回忆定义3, 域$\mathfrak{K}$上n维空间V 称为是对于线性算子$\mathcal{A}$: $V \to V$ 是循环的, 如果在V 中有一个向量$\mathbf{v}$, 使得

$$V = \langle \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n \rangle,$$

其中

$$\mathbf{e}_i = \mathcal{A}^{n-i}\mathbf{v}, \quad i = 1, 2, \cdots, n.$$

此时, 也把$\mathbf{v}$ 称为是**循环的向量**. 因为$(\mathbf{e}_i)$ 是空间V 的基底, 所以

$$\mathcal{A}^n\mathbf{v} = \sum_{i=0}^{n-1} \alpha_i \mathcal{A}^i \mathbf{v},$$

而且系数$\alpha_i \in \mathfrak{K}$是唯一确定的. 所以, 线性算子$\mathcal{A}$ 在这个基底之下对应的就是被称为**循环块的矩阵**, 形如

$$A = \begin{pmatrix} \alpha_{n-1} & 1 & 0 & \cdots & 0 & 0 \\ \alpha_{n-2} & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_1 & 0 & 0 & \cdots & 0 & 1 \\ \alpha_0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{12}$$

相反: 如果算子$\mathcal{A}$ 在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ 之下的矩阵是个循环块, 那么, $\mathbf{v} = \mathbf{e}_n$ 就是循环的向量, 而且$\mathbf{e}_i = \mathcal{A}^{n-i}\mathbf{e}_n$ (对i往下面归纳).

我们指出, $\mathcal{A}$ 对应的循环块的形状与初始的循环向量的选择没有关系. 之所以可以充分相信这一点, 在于块(12)的第1列是由$\mathcal{A}$ 的极小多项式, 即

$$\mu_{\mathcal{A}}(t) = f(t) = t^n - \sum_{i=0}^{n-1} \alpha_i t^i$$

的系数组成的.

事实上, $f(\mathcal{A}) = \mathcal{O}$, 从而

$$f(\mathcal{A})[\mathcal{A}^i\mathbf{v}] = \mathcal{A}^i[f(\mathcal{A})\mathbf{v}] = \mathbf{0},$$

而诸向量$\mathcal{A}^i\mathbf{v}$ 生成V . 另一方面, 对于任意次数$< n$ 的多项式$g(t)$, $g(\mathcal{A}) \neq \mathcal{O}$, 因为, 若不然, 可把算子$g(\mathcal{A}) = \mathcal{O}$ 作用到循环向量$\mathbf{v}$ 上, 我们就得到基底向量$\mathcal{A}^i\mathbf{v}$ 之间的一个非平凡的线性关系式.

ii) 空间的循环度. 对照前面的研究, 如果空间V 对于线性算子$\mathcal{A}$ 是循环的, 那么, 它的维数等于算子$\mathcal{A}$ 的极小多项式的次数. 可见, 极小多项式与特征多项式是一致的. 反之亦然(见§3的习题4).

iii) *任意线性算子在适当的基底之下的矩阵可以表成循环块的直和.* 这个证明可归结为类似于有关若尔当标准型定理的证明. 我们应该研究特征多项式的因子$p_i(t)^{r_i}$, 其中$p_i(t)$ 是域$\mathfrak{K}$上的不可约多项式, 来代替原来特征多项式的因子$(t-\lambda_i)^{n_i}$. 如果限制条件, 在所有循环块的极小多项式都不可约的时候, 唯一性定理同样也是成立. 没有这个限制条件, 唯一性定理就不对了: 循环的空间可以是两个极小多项式互素的子空间的直和.

习　　题

1. 运用例4的矩阵S 和对$J(S)$ 的表达式, 计算出矩阵

$$A = \begin{pmatrix} m & -1 & \cdots & -1 \\ -1 & m & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & \cdots & m \end{pmatrix}$$

的行列式, 并把它写成det $A = \chi_S(m+1)$ 形式.

2. 精确到相似, 下列矩阵

$$A_1 = J_2(0)\dot{+}J_1(0)\dot{+}J_1(0), \quad A_2 = J_2(0)\dot{+}J_2(0),$$
$$A_3 = J_3(0)\dot{+}J_1(0), \quad A_4 = J_4(0).$$

就穷尽了所有的非零的4×4阶的幂零矩阵. 矩阵

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

显然都是幂零的, 它们各相似于哪个A_i ?

3. (1) 已知道$\chi_{\mathcal{A}}(t) = (t-3)^4(t+2)$ 且rank$(\mathcal{A} - 3\mathcal{E}) = 2$, 求出$J(\mathcal{A})$;

(2) 在rank$(\mathcal{A} - 3\mathcal{E}) = 1, 3, 4$ 的情形, 能唯一地复原$J(\mathcal{A})$ 吗?

4. (1) 指明, 矩阵

$$A=\begin{pmatrix}6&2&-2\\-2&2&2\\2&2&2\end{pmatrix},\quad B=\begin{pmatrix}6&2&2\\-2&2&0\\0&0&2\end{pmatrix}$$

有相同的特征多项式;

(2) 求出$\mu_A(t)$ 和$\mu_B(t)$;

(3) 求出$J(A)$ 和$J(B)$.

5. 设域$\mathfrak{K}$的特征数为零, A 是$\mathfrak{K}$上$n\times n$ 矩阵. 证明, A 只有在$\mathrm{tr}(A^k)=0$, $1\leqslant k\leqslant n$ 时, 才能幂零.

6. 证明, 矩阵$A\in M_n(\mathbb{C})$与${}^tA$ 恒共轭.

7. 对于矩阵$A\in M_n(\mathbb{C})$, 关系式$A^N=E$ 成立, 当且仅当, A 可以对角化同时它的特征值都是N 次单位根.

8. 直接验证, 矩阵

$$A=\begin{pmatrix}1&2&0\\0&1&2\\2&0&1\end{pmatrix}\in \mathrm{Mag}_3(\mathbb{Q}),$$

但$A^2\not\in \mathrm{Mag}_3(\mathbb{Q})$, 也就是说, 不同于半幻方(见第1章§2习题10), 幻方不构成一个环. 更令人感到意外的是下面的命题:

如果$A\in \mathrm{Mag}_3(\mathbb{Q})$, 那么, 对任意奇数$m\geqslant 1$ 都有$A^m\in \mathrm{Mag}_3(\mathbb{Q})$.

根据哈密顿–凯莱定理证明这一论断.

9. 验证, 对任意$m\geqslant 2$, 矩阵A^m 都不是一个幻方, 其中

$$A=\begin{pmatrix}2&0&0&0\\0&0&1&1\\0&1&0&1\\0&1&1&0\end{pmatrix}\in \mathrm{Mag}_4(\mathbb{Q}).$$

利用矩阵A , 证明, 对所有的$n\geqslant 4$, 都有一个$n\times n$ 阶的幻方矩阵, 而它的m次幂($m\geqslant 2$)却不是幻方矩阵.

10. 把矩阵$A=J_1(\lambda)\dot{+}J_2(\mu)$, $\lambda\neq\mu$ 写成$A=S+N$ 的样子, 其中$S=\mathrm{diag}(\lambda,\mu,\mu)$. 更详细地:

$$A=\begin{pmatrix}\lambda&0&0\\0&\mu&1\\0&0&\mu\end{pmatrix},S=\begin{pmatrix}\lambda&0&0\\0&\mu&0\\0&0&\mu\end{pmatrix},N=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix}.$$

把S 和N 写成A 的多项式$s(A)$ 和$n(A)$ 形式(按照哈密顿–凯莱定理, 可计算出, $\deg\ s(t)\leqslant 2$, $\deg\ n(t)\leqslant 2$; 照惯例, $A^0=E$).

11. 设V 是个复的n维空间, $\mathcal{A}$ 是V 上的线性算子. 证明, $\mathcal{A}$ 允许唯一地表达成

$$\mathcal{A}=\mathcal{S}+\mathcal{N},\quad \mathcal{S}\mathcal{N}=\mathcal{N}\mathcal{S},$$

其中$\mathcal{S}$是可对角化的, 而$\mathcal{N}$是一个幂零线性算子, 同时, $\mathcal{S}$和$\mathcal{N}$都可以表达成$\mathcal{A}$的多项式形式($\mathcal{S}$被称为**半单的**, 而$\mathcal{N}$被称为是算子$\mathcal{A}$的**复幂零算子**).

12. 对任意自然数k, 计算出$(J_n(\lambda))^k$.

13. 证明, 对任意三个矩阵$X, Y, Z \in M_2(\mathfrak{K})$, 必有$\left[[X,Y]^2, Z\right] = 0$, 这里的$[X,Y] = XY - YX$是矩阵$X, Y$的换位子.

第3章　带有纯量乘积的向量空间

在第1章, 我们阐述了双线性型理论, 某种程度上, 就是为了可以把一般的线性空间转换到内容更加丰富多彩, 甚至可以说是更为熟悉的结构——度量空间. 回想一下, 我们在3维几何中的丰富成果, 很大程度上, 是以向量代数的两个补充概念为前提的, 它们是向量的长度和两个向量之间的夹角. 把与基础域关系不大的线性的纯粹定性性质转化到向量空间对象之间的定量的关系, 要求我们把精力实质上集中到两个纯量域$\mathbb{R}$ 和$\mathbb{C}$ 上. 由于本身的重要性, 也由于研究新的形式的需要, 复向量的几何得到了特殊的讨论.

§1　欧几里得向量空间

1. 直观理解与定义　两个向量的纯量积在空间$\mathbb{R}^2$ 和$\mathbb{R}^3$ 的解析几何中起了重要的作用. 这种纯量积是作为这些向量的长度以及它们之间的夹角的余弦的乘积引进来的. 在给定的直角坐标系, 向量$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_3\mathbf{e}_3$ 的长度$\|\mathbf{x}\|$ 按公式

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

定义. 所以, 对任意$\mathbf{x} \neq \mathbf{0}$ 都有$\|\mathbf{x}\| > 0$. 长度的平方$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + x_3^2$ 可以解释成正定二次型的值. 与其配极的对称的双线性型$(*|*)$, 用到任意两个向量$\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ 上, 即可与数(或纯量)$(\mathbf{x}|\mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3$ 相提并论. 设$\varphi = (\widehat{\mathbf{x}, \mathbf{y}})$ 是向量$\mathbf{x}$和$\mathbf{y}$ 之间的夹角, $\mathbf{z} = \mathbf{y} - \mathbf{x}$ 是它们的差(图3). 把通常初等几何学中的余弦公式应用到以$\mathbf{x}$, $\mathbf{y}, \mathbf{z}$为边的三角形上, 可以导出

$$\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\,\|\mathbf{x}\|\,\|\mathbf{y}\| \cos\varphi.$$

另一方面, 利用型$(*|*)$ 的双线性性和对称性, 我们可以得到$\|\mathbf{z}\|^2 = \|\mathbf{y} - \mathbf{x}\|^2 = (\mathbf{y} -$

$\mathbf{x}|\mathbf{y}-\mathbf{x}) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2(\mathbf{x}|\mathbf{y})$. 比较$\|\mathbf{z}\|^2$ 的两个表达式, 得到

$$(\mathbf{x}|\mathbf{y}) = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\varphi,$$

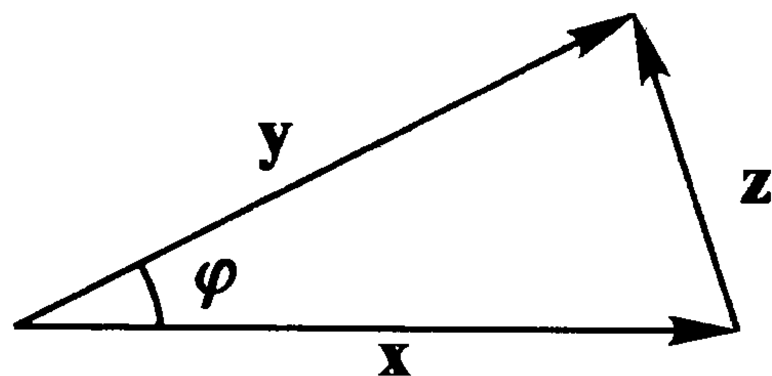


图3

即纯量$(\mathbf{x}|\mathbf{y})$ 与通常的向量$\mathbf{x}$, $\mathbf{y}$ 的纯量乘积一致. 这种观察提示一种合理的方式来引进$\mathbb{R}^n$ 中向量间的纯量乘积:

$$(\mathbf{x}|\mathbf{y}) = \sum_{i=1}^{n} x_i y_i. \tag{1}$$

但是, 这个定义让人感到包含某些个人意愿, 即与基底的选择有关系, 为了排除这一点, 我们引入下面一般的

定义1 实的向量空间V 带有一个指定的对称的双线性型$(\mathbf{x},\mathbf{y}) \mapsto (\mathbf{x}|\mathbf{y})$, 它对应的二次型$\mathbf{x} \mapsto (\mathbf{x}|\mathbf{x})$ (或简单地, $(\mathbf{x}|\mathbf{x})$)是正定的, 就称空间V 是一个**欧几里得向量空间**.

在一般情况下, 对称双线性型$(*|*)$ 在向量$\mathbf{x},\mathbf{y} \in V$ 处的值$(\mathbf{x}|\mathbf{y})$ 被称为是它们的**纯量积**. 我们用符号 $(\mathbf{x}|\mathbf{y})$ 代替通常的$f(\mathbf{x},\mathbf{y})$. 有责任强调这样一个事实, 就是在二次型的无数的集合中分出一个来使它成为定义欧几里得空间的基础. 因为我们准备借助$(\mathbf{x}|\mathbf{y})$ 导入长度和夹角的概念, 所以, 可将在这种并非单值情形获得的东西与度量直线上的线段时比例的选择的随意性相比较. 经常把纯量乘积表达成$(\mathbf{x},\mathbf{y})$ 或者$\langle\mathbf{x},\mathbf{y}\rangle$. 但是, 我们已经有简单的向量对$(\mathbf{x},\mathbf{y})$ (笛卡儿积$V \times V$ 的元素), 而$\langle\mathbf{x},\mathbf{y}\rangle$ 又是$\mathbf{x}$, $\mathbf{y}$ 生成的子空间. 将来, 我们把双线性型$(*|*)$ 作为空间$\mathcal{L}_2(V,\mathbb{R})$ 和它在任意向量$\mathbf{x}$, $\mathbf{y}$ 处的值$(\mathbf{x}|\mathbf{y})$ 等同起来.

这就是说, 按定义1, 欧几里得空间是一个对$(V\ ,\ (*|*)\ )$, 其中V 是$\mathbb{R}$ 上的向量空间, 而$(*|*)$是V 上一个固定的对称的双线性型. 我们再一次把纯量乘积的基本性质列出来:

i) $(\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x}), \quad \forall \mathbf{x},\mathbf{y} \in V$;

ii) $(\alpha\mathbf{x} + \beta\mathbf{y}|\mathbf{z}) = \alpha(\mathbf{x}|\mathbf{z}) + \beta(\mathbf{y}|\mathbf{z}), \quad \forall \alpha,\beta \in \mathbb{R}$;

iii) $(\mathbf{x}|\mathbf{x}) > 0, \quad \forall \mathbf{x} \neq \mathbf{0}\ ((\mathbf{0}|\mathbf{x}) = 0)$.

关系式(1)所给的纯量乘积(称它为**标准纯量积**), 显然满足这些性质且适合一般定义. 换言之, 后者的意义更广泛.

例1 设$V = P_n$ 是次数$\leqslant n-1$ 的实多项式的向量空间, 对任意两个向量(多项式)$f,g \in V$, 数

$$(f|g) = \int_a^b f(t)g(t)dt \tag{2}$$

($[a,b]$ 是$\mathbb{R}$ 上一个固定的区间)同样给出V上的一个向量间的纯量乘积, 这很容易从定积分的性质看出来. 并没有把纯量乘积充分地用自然基底$1, t, \cdots, t^{n-1}$ 的术语表达. 应当注意, 纯量乘积是用关系式(2)在连续函数(在区间$[a,b]$ 上)的无穷维向量空间$C(a,b)$ 上给出的. 相对应的无穷维欧几里得空间用符号$C_2(a,b)$ 代表.

2. 基本的度量概念 设V 是个带有纯量乘积$(\mathbf{x}|\mathbf{y})$ 的欧几里得空间.

定义2 称非负实数

$$\|\mathbf{v}\| = \sqrt{(\mathbf{v}|\mathbf{v})}. \tag{3}$$

是任意向量$\mathbf{v} \in V$的**长度**或**模**

因为$(\mathbf{v}|\mathbf{v}) \geqslant 0$, 所以任意向量$\mathbf{v}$ 的长度都是完全确定的. 而且$\mathbf{v} \neq \mathbf{0} \Rightarrow \|\mathbf{v}\| > 0$. 如果$\lambda \in \mathbb{R}$; 那么$\|\lambda\mathbf{v}\| = \sqrt{(\lambda\mathbf{v}|\lambda\mathbf{v})} = |\lambda| \cdot \|\mathbf{v}\|$.

在此处, 我们要注意, 欧几里得向量空间V 的任意一个子空间U 本身也是个欧几里得向量空间, 因为纯量乘积$(\mathbf{x}|\mathbf{y})$ 在U上导出的限制定义了$U \times U \to \mathbb{R}$ 的双线性函数. 显然, 它保存了对称性和正定性. 特别地, 域$\mathbb{R}$ 本身可以被看成是个1维的实的向量空间, 在它上面, 向量的长度和通常实数的绝对值一致. 在一般情形, 我们用$|*|$和$\|*\|$ 区分它们.

称长度为1的向量是**标准的**. 任意向量$\mathbf{x} \neq \mathbf{0}$, 把它乘上适当的数, 就可以标准化了. 即, 对$\mathbf{x}' = \dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$. 我们有

$$\|\mathbf{x}'\| = \left\|\frac{1}{\|\mathbf{x}\|}\mathbf{x}\right\| = \frac{1}{\|\mathbf{x}\|}\|\mathbf{x}\| = 1.$$

在进入两个向量之间的夹角之前, 我们再次转向纯量乘积的性质iii).

定理1(柯西–布尼亚科夫斯基不等式) 欧几里得空间中, 对任意向量$\mathbf{x}, \mathbf{y}$, 不等式

$$|(\mathbf{x}|\mathbf{y})| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\| \tag{4}$$

都成立.

证明 由纯量乘积的正定性(性质iii))得到

$$\lambda^2(\mathbf{x}|\mathbf{x}) - 2\lambda(\mathbf{x}|\mathbf{y}) + (\mathbf{y}|\mathbf{y}) = (\lambda\mathbf{x} - \mathbf{y}|\lambda\mathbf{x} - \mathbf{y}) \geqslant 0, \tag{5}$$

其中λ是任意实数.

固定向量$\mathbf{x}, \mathbf{y} \in V$, 我们把(5)的左侧看成是二次三项式f, 因为对任意数$\lambda \in \mathbb{R}$ 都有$f(\lambda) \geqslant 0$, 所以, 它的判别式$D(f) = (2(\mathbf{x}|\mathbf{y}))^2 - 4(\mathbf{x}|\mathbf{x})(\mathbf{y}|\mathbf{y})$ 应当满足不等式$D(f) \leqslant 0$, 从而有

$$(\mathbf{x}|\mathbf{y})^2 \leqslant (\mathbf{x}|\mathbf{x}) \cdot (\mathbf{y}|\mathbf{y}). \tag{6}$$

把(6)两侧的非负的平方数和(3)的向量长度所用的定义联系到一起, 我们就得到不等式(4), 其左侧是个纯量$(\mathbf{x}|\mathbf{y})$的绝对值. □

说明 如果$|(\mathbf{x}|\mathbf{y})| = \|\mathbf{x}\|\,\|\mathbf{y}\|$，那么$D(f)=0$, 即二次三项式$f$只有一个根$\lambda_0$. 对照(5)有$(\lambda_0\mathbf{x}-\mathbf{y}|\lambda_0\mathbf{x}-\mathbf{y})=0$，从而$\mathbf{y}=\lambda_0\mathbf{x}$. 所以, 只有是共线的(成比例的)向量, 它们的纯量积的绝对值才能等于它们的长度的乘积.

例2 应用到标准纯量乘积(1)和空间$C_2(a,b)$ 上的纯量积上, 不等式(4)就变成

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2} \tag{7}$$

和相应的

$$\left|\int_a^b f(t)g(t)dt\right| \leqslant \sqrt{\int_a^b f^2(t)dt} \cdot \sqrt{\int_a^b g^2(t)dt}. \tag{8}$$

不等式(8)在数学分析中起着重要作用.

柯西–布尼亚科夫斯基不等式意味着

$$-1 \leqslant \frac{(\mathbf{x}|\mathbf{y})}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|} \leqslant 1.$$

由此可见, 关系式$(\mathbf{x}|\mathbf{y})/\,(\|\mathbf{x}\|\cdot\|\mathbf{y}\|)$ 是一个完全确定的角φ 的余弦:

$$\cos\varphi = \frac{(\mathbf{x}|\mathbf{y})}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}, \quad 0 \leqslant \varphi \leqslant \pi. \tag{9}$$

也就是说, 这个角φ，按定义, 可以看成是向量$\mathbf{x}$和$\mathbf{y}$之间的夹角.

定义3 *如果向量$\mathbf{x}$, $\mathbf{y}$之间的夹角为$\pi/2$, 也就是$(\mathbf{x}|\mathbf{y})=0$, 那么, 就说它们是***正交的***(表示成$\mathbf{x}\perp\mathbf{y}$).*

零向量正交于每个向量$\mathbf{x}\in V$. 还应注意

$$\mathbf{x}\perp\mathbf{y} \Rightarrow \|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

(毕达哥拉斯定理), 而且这个初等几何命题是纯量乘积性质公式的一个推论. 关于两两正交的向量$\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\cdots$的命题, 更一般些, 是

$$\|\mathbf{x}+\mathbf{y}+\mathbf{z}+\mathbf{u}+\cdots\|^2 = \|\mathbf{x}\|^2+\|\mathbf{y}\|^2+\|\mathbf{z}\|^2+\|\mathbf{u}\|^2+\cdots.$$

命题的实质是验证, 恒有

$$\|\mathbf{x}\| = \|\mathbf{y}\| \Rightarrow (\mathbf{x}+\mathbf{y})\perp(\mathbf{x}-\mathbf{y})$$

(菱形的两个对角线相交成直角).

由定理1可以推出

推论(三角不等式) *向量$\mathbf{x}$, $\mathbf{y}$与$\mathbf{x}+\mathbf{y}$的长度可用不等式*

$$\|\mathbf{x}\pm\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \tag{10}$$

联系在一起.

证明 实际上, 用不等式(4), 可得到

$$\begin{aligned}\|\mathbf{x}\pm\mathbf{y}\|^2 &= \|\mathbf{x}\|^2+\|\mathbf{y}\|^2\pm 2(\mathbf{x}|\mathbf{y})\leqslant \|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2|(\mathbf{x}|\mathbf{y})|\\ &\leqslant \|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2\,\|\mathbf{x}\|\cdot\|\mathbf{y}\|=(\|\mathbf{x}\|+\|\mathbf{y}\|)^2.\end{aligned}\qquad\square$$

例3 在函数空间$C_2(a,b)$ 中不等式(10)变成形如

$$\sqrt{\int_a^b (f(t)\pm g(t))^2dt}\leqslant\sqrt{\int_a^b f(t)^2dt}+\sqrt{\int_a^b g(t)^2dt}$$

(闵可夫斯基不等式).

3. 正交化过程 在带有纯量乘积的标准向量空间$\mathbb{R}^n$中, 向量$\mathbf{e}_i=(0,\cdots,1,\cdots 0)$, $i=1,\cdots,n$ 两两正交而且作成一个基底. 自然地期待, 在任意的欧几里得向量空间V中都能选出具有类似性质的基底. 为准确地说明这一点, 我们要有

定义4 欧几里得向量空间V 的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 被称为是**正交的**, 如果, 当$i\neq j$时, $(\mathbf{e}_i|\mathbf{e}_j)=0$, $i,j=1,2,\cdots,n$. 进一步, 如果当$j=1,2,\cdots,n$ 时总有$(\mathbf{e}_j|\mathbf{e}_j)=1$, 那么, 称这个基底是个**标准正交基底**(或者**规格化的正交基底**).

换言之, 在标准正交基底中所有的向量$\mathbf{e}_i$ 都具有单位长度. 由每个正交基底都可以得到一个标准正交基底, 只要将每个向量$\mathbf{e}_i$ 标准化.

我们指出下面的几乎是明显的事实.

定理2 任意非零的相互正交的向量$\mathbf{e}_1,\cdots,\mathbf{e}_m\in V$ 都是线性无关的. 如果, 在此基础上还有$\dim\ V=n$ 同时$m=n$, 那么, 向量组$\mathbf{e}_i$ 在V 中构成一个标准正交基底.

证明 第二个断言可由第一个断言推出来; 现在就来证明第一个. 设

$$\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\cdots+\alpha_m\mathbf{e}_m=\mathbf{0}$$

是向量$\mathbf{e}_1,\cdots,\mathbf{e}_m$ 之间的一个非平凡的线性关系式. 再设, 比如, $\alpha_k\neq 0$. 用$\mathbf{e}_k$ 在该等式的两侧做纯量乘积, 得

$$\begin{aligned}0&=(\mathbf{0}|\mathbf{e}_k)=(\alpha_1\mathbf{e}_1+\cdots+\alpha_m\mathbf{e}_m|\mathbf{e}_k)\\ &=\alpha_1(\mathbf{e}_1|\mathbf{e}_k)+\cdots+\alpha_k(\mathbf{e}_k|\mathbf{e}_k)+\cdots+\alpha_m(\mathbf{e}_m|\mathbf{e}_k)=\alpha_k(\mathbf{e}_k|\mathbf{e}_k),\end{aligned}$$

因为, 按条件$i\neq k$ 时, $(\mathbf{e}_i|\mathbf{e}_k)=0$. 另一方面, $(\mathbf{e}_k|\mathbf{e}_k)\neq 0$, 于是, 我们推得结论, $\alpha_k=0$. 得到的矛盾证明了定理. $\square$

定理2的结论要在已有一个固定的正交基底的结构时才可以应用, 但应该及早建立它的存在性.

定理3 在任意n维欧几里得空间V 中都必存在一个标准正交基底.

证明 V 上的二次型q:

$$q(\mathbf{x})=(\mathbf{x}|\mathbf{x})=\|\mathbf{x}\|^2$$

是正定的. 所以对于q, 像对所有其他正定型一样(见第1章§4第8目), 存在V 的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$, 在这个基底之下, q可以表成标准型

$$q(\mathbf{x}) = x_1^2 + x_2^2 + \cdots + x_n^2$$

$(\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n)$. 任意向量$\mathbf{x}$和$\mathbf{y}$ 的纯量乘积就是(1). 但是, 这就意味着$(\mathbf{e}_i|\mathbf{e}_j) = \delta_{ij}$, 也就是说, $(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ 是一个标准正交基底(在给定的情形, 更准确地说, 规格化正交基底). □

注意, 在标准正交基底之下, 向量$\mathbf{x}$的坐标等于用相应的基底向量与$\mathbf{x}$ 作的纯量乘积:

$$(\mathbf{x}|\mathbf{e}_i) = x_i \tag{10'}$$

定义5 称纯量乘积$(\mathbf{x}|\mathbf{e})$ 是向量$\mathbf{x}$ 在直线$\langle\mathbf{e}\rangle_{\mathbb{R}}$ 上的**投影**, 其中$\mathbf{e}$是个长度为1的向量.

这样一来, 我们可以说, 向量$\mathbf{x}$在标准正交基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ 之下的坐标与$\mathbf{x}$在坐标轴$\langle\mathbf{e}_i\rangle_{\mathbb{R}}$ 上的投影是一致的. 借助于通常称之为**格拉姆–施密特正交化过程**, 可以实现标准正交基底的具体构作, 这个方法在分析学和几何学的很多不同的问题中都会遇到. 我们先要注意到, 与给定的向量$\mathbf{v}$正交的所有向量的集合是一个子空间, 可称为是$\mathbf{v}$ 的**正交补**. 事实上, 如果$\mathbf{x}\perp\mathbf{v}, \mathbf{y}\perp\mathbf{v}$, 即$(\mathbf{x}|\mathbf{v}) = (\mathbf{y}|\mathbf{v}) = 0$, 那么, 就有

$$(\alpha\mathbf{x} + \beta\mathbf{y}|\mathbf{v}) = \alpha(\mathbf{x}|\mathbf{v}) + \beta(\mathbf{y}|\mathbf{v}) = 0, \forall\alpha, \beta \in \mathbb{R}.$$

同样可以说, 向量$\mathbf{v}$正交于子空间$U \subset V$, 如果$\mathbf{v}\perp\mathbf{u}$, $\forall\mathbf{u} \in U$. 显然, $\mathbf{v}\perp U \Leftrightarrow \mathbf{v}\perp\mathbf{e}_i$, $i = 1, 2, \cdots, m$, 其中, $\mathbf{e}_1, \cdots, \mathbf{e}_m$ 是U 的基底的向量.

最后, 我们引入

定义6 所有的正交于子空间$U \subset V$ 的向量$\mathbf{x} \in V$ 的集合是一个子空间(由于条件$\mathbf{x}\perp U$ 的线性性质), 记为$U^{\perp}$, 称它是U的**正交补**.

定理4 (正交化过程) 设$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 是n维欧几里得向量空间V 中的一组m个线性无关的向量. 那么, 存在一个正交向量组$\mathbf{e}'_1, \cdots, \mathbf{e}'_m$ 使得线性包络$L_i = \langle\mathbf{e}_1, \cdots, \mathbf{e}_i\rangle$ 和$L'_i = \langle\mathbf{e}'_1, \cdots, \mathbf{e}'_i\rangle$ 当$i = 1, 2, \cdots, m$ 时都重合, $m \leqslant n$.

证明 把$\mathbf{e}'_1$ 取成$\lambda\mathbf{e}_1$, 其中$\lambda = \|\mathbf{e}_1\|^{-1}$. 因为, $L_1 = \langle\mathbf{e}_1\rangle = \langle\mathbf{e}'_1\rangle = L'_1$, 这就给出了命题当$m = 1$ 的情形. 设, 当$1 \leqslant k < m$ 的情形, 已经构建出了所需要的向量组$\mathbf{e}'_1, \cdots, \mathbf{e}'_k$ $(L_i = L'_i, i = 1, 2, \cdots, k)$. 我们来指明如何找出$\mathbf{e}'_{k+1}$.

向量$\mathbf{e}_{k+1}$ 不包含在$L_k = L'_k$ 中(否则, $\mathbf{e}_{k+1}$ 可用$\mathbf{e}_1, \cdots, \mathbf{e}_k$ 线性表示), 所以, $L_{k+1} = \langle\mathbf{e}_1, \cdots, \mathbf{e}_k, \mathbf{v}\rangle$, 其中,

$$\mathbf{v} = \mathbf{e}_{k+1} - \sum_{i=1}^{k} \lambda_i\mathbf{e}'_i,$$

而$\lambda_1, \cdots, \lambda_k$ 是任意纯量. 我们力求选好λ_i 使得$\mathbf{v}$ 与L'_k 正交. 要做到这一点, 正如我们已经知道的, 充分而且必要地满足条件

$$0 = (\mathbf{v}|\mathbf{e}'_j) = (\mathbf{e}_{k+1}|\mathbf{e}'_j) - \left(\sum_{i=1}^{k} \lambda_i \mathbf{e}'_i|\mathbf{e}'_j\right)$$
$$= (\mathbf{e}_{k+1}|\mathbf{e}'_j) - \sum_{i=1}^{k} \lambda_i(\mathbf{e}'_i|\mathbf{e}'_j) = (\mathbf{e}_{k+1}|\mathbf{e}'_j) - \lambda_j, \quad j = 1, \cdots, k.$$

这样一来, 取$\lambda_j = (\mathbf{e}_{k+1}|\mathbf{e}'_j)$ 就得到了$\mathbf{v} \neq \mathbf{0}$ 正交于L'_k . 设$\mathbf{e}'_{k+1} = \mu\mathbf{v}$, $\mu = \|\mathbf{v}\|^{-1}$, 我们就得到了标准正交组$\mathbf{e}'_1, \cdots, \mathbf{e}'_{k+1}$, 而且$L_{k+1} = L'_{k+1}$. 最后, 就得到所要求的向量组$\mathbf{e}'_1, \cdots, \mathbf{e}'_m$. □

推论　*欧几里得向量空间V 中所有的标准正交向量组都可以扩充成标准正交基底.*

证明　按照第1章§2的定理3, 据条件, 已有的标准正交组$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 可以扩充成基底$\mathbf{e}_1, \cdots, \mathbf{e}_m$, $\mathbf{e}_{m+1}, \cdots, \mathbf{e}_n$. 对于这个基底, 利用定理4中论述的正交化过程, 可以不触及前面的m个向量. □

运用一个很接近前述定理的正交过程的手法可以证明下面的论断.

定理5　*设L 是有限维欧几里得向量空间V 的一个子空间, $L^\perp$ 是它的正交补, 那么,*

$$V = L \oplus L^\perp, \quad L^{\perp\perp} = L. \tag{11}$$

证明　我们在L 中取任意一个标准正交基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m)$. 设$\mathbf{w} \in V$. 知向量

$$\mathbf{v} = \mathbf{w} - \sum_{i=1}^{m} (\mathbf{w}|\mathbf{e}_i)\mathbf{e}_i.$$

因为$(\mathbf{v}|\mathbf{e}_j) = (\mathbf{w}|\mathbf{e}_j) - \sum_{i=1}^{m} (\mathbf{w}|\mathbf{e}_i)(\mathbf{e}_i|\mathbf{e}_j) = (\mathbf{w}|\mathbf{e}_j) - (\mathbf{w}|\mathbf{e}_j) - (\mathbf{w}|\mathbf{e}_j) \cdot 1 = 0; j = 1, 2, \cdots, m$, 所以, 向量$\mathbf{v}$ 正交于子空间L. 这意味着, $\mathbf{w} = \mathbf{u} + \mathbf{v}$, 其中$\mathbf{u} = \sum_{i=1}^{m} (\mathbf{w}|\mathbf{e}_i)\mathbf{e}_i \in L$ 而$\mathbf{v} \in L^\perp$. 也就是, $V = L + L^\perp$.

设$\mathbf{x} \in L \cap L^\perp$, 因为$\mathbf{x} \in L$, 所以$(\mathbf{x}|L) = 0$. 但是, 特别地, $\mathbf{x} \in L^\perp$, 所以$(\mathbf{x}|\mathbf{x}) = 0$, 从而$\mathbf{x}=\mathbf{0}$. 从而, $V = L \oplus L^\perp$ 是直和.

由分解式$\mathbf{w} = \mathbf{u} + \mathbf{v}$ $(\mathbf{u} \in L, \mathbf{v} \in L^\perp)$ 得到$(\mathbf{w}|\mathbf{u}) = (\mathbf{u} + \mathbf{v}|\mathbf{u}) = (\mathbf{u}|\mathbf{u}) + (\mathbf{v}|\mathbf{u}) = \|\mathbf{u}\|^2$, 类似地, $(\mathbf{w}|\mathbf{v}) = \|\mathbf{v}\|^2$. 现在, 如果$\mathbf{w} \in L^{\perp\perp}$, 那么, $(\mathbf{w}|\mathbf{v}) = 0$ 且$\|\mathbf{v}\|^2 = 0$, 导致$\mathbf{w} = \mathbf{u} \in L$. 可见, $L^{\perp\perp} \subseteq L$. 其次, 因为$L^{\perp\perp} = (L^\perp)^\perp$ 是正交于$L^\perp$ 的子空间, 而$(L|L^\perp) = 0$, 故$L \subseteq L^{\perp\perp}$. 从而$L^{\perp\perp} = L$. □

4. 欧几里得向量空间的同构　我们看到了, 在欧几里得向量空间中选择一个标准正交基底就给出了把纯量乘积$(\mathbf{x}|\mathbf{y})$记成标准型(1)的可能性. 这个事实实际上意味着向量空间V 和$\mathbb{R}^n$就其自己度量性质来说是没有差别的. 更精确的命题可以表述成

定理6 任意两个维数相同的欧几里得向量空间V，V'都是同构的. 意思就是，存在向量空间的同构映射$f: V \to V'$(见第1章§2第3目中的定义)，它还保持纯量乘积，即

$$(\mathbf{x}|\mathbf{y}) = (f(\mathbf{x})|f(\mathbf{y}))' \tag{12}$$

($(*|*)'$是V'上的纯量乘积).

证明 研究V的一个标准正交基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$与$V'$的某个标准正交基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$. 对应

$$f: \mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n \mapsto \mathbf{x}' = x_1\mathbf{e}'_1 + \cdots + x_n\mathbf{e}'_n,$$

显然是个双射. 如同在第1章§2第3目的定理5的情况一样，可以直接确认，f是向量空间的同构映射. 因为在V和V'中计算纯量乘积$(\mathbf{x}|\mathbf{y})$和$(\mathbf{x}'|\mathbf{y}')'$都是按着同一个公式(1)进行(鉴于基底的选择)，所以，欧几里得向量空间同构的条件(12)同样得到了满足. □

刚刚证明的这个定理允许用V中向量作用与V上纯量乘积术语描述的任何命题转换成几何语言. 反过来，对于空间$\mathbb{R}^2$和$\mathbb{R}^3$中的度量定理应该在所有的维数$\leqslant 3$的欧几里得向量空间V上保持正确.

还有一种涉及同构的说法，即按第1章§3的思想研究与欧几里得向量空间V共轭(对偶)的空间V^*.

显然，在固定任意一个向量$\mathbf{v} \in V$的情况下，映射$\mathbf{\Phi}: \mathbf{x} \mapsto (\mathbf{v}|\mathbf{x})$决定了一个线性型:

$$\mathbf{\Phi}_\mathbf{v} = (\mathbf{v}|*): V \to \mathbb{R},$$

也就是说，$(\mathbf{v}|*) \in V^*$.

定理7 映射$\mathbf{\Phi}: \mathbf{v} \mapsto (\mathbf{v}|*) = \mathbf{\Phi}_\mathbf{v}$是向量空间$V$到$V^*$的自然同构. 在这个同构之下，欧几里得向量空间$V$的标准正交基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$与$V^*$中与其对偶的基底$f_1, \cdots, f_n$相匹配.

证明 因为纯量乘积$(\mathbf{v}|\mathbf{x})$对$\mathbf{v}$是线性的，所以映射$\mathbf{\Phi}$也是线性的:

$$\mathbf{\Phi}_{(\alpha\mathbf{u}+\beta\mathbf{v})} = (\alpha\mathbf{u}+\beta\mathbf{v}|*) = \alpha(\mathbf{u}|*) + \beta(\mathbf{v}|*) = \alpha\mathbf{\Phi}_\mathbf{u} + \beta\mathbf{\Phi}_\mathbf{v}.$$

其次，$\operatorname{Ker} \mathbf{\Phi} = 0$，因为$\mathbf{v} \in \operatorname{Ker} \mathbf{\Phi} \Rightarrow (\mathbf{v}|\mathbf{x}) = 0, \forall \mathbf{x} \in V$，且，特别地，$(\mathbf{v}|\mathbf{v}) = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

作为空间V^*的任意元素，线性型$(\mathbf{v}|*)$一定可以由向量空间V^*的对偶于基底$(\mathbf{e}_i)$的基底$e^1, \cdots, e^n \in V^*$线性表示. 特别地

$$\mathbf{\Phi}_{\mathbf{e}_i} = (\mathbf{e}_i|*) = \sum_{j=1}^{n} a_{ij}e^j, \quad i = 1, \cdots, n.$$

因为$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 是标准正交基底, 所以

$$a_{ij}=\sum_{k=1}^{n}a_{ik}\delta_{jk}=\sum_{k=1}^{n}a_{ik}e^k(\mathbf{e}_j)=(\mathbf{e}_i|\mathbf{e}_j)=\delta_{ij},$$

于是有

$$(\mathbf{e}_i|*)=\mathbf{e}^i. \tag{13}$$

这就给出了Φ 的满射性, 进而, 给出了Φ 的双射性. 与此同时, 关系式(13)建立了定理断言的结论的正确性.

这就是说, 在欧几里得向量空间V 中每个向量$\mathbf{v}\in V$ 都可以看成是一个线性型 $\mathbf{v}:V\to\mathbb{R}$. 在这样的等同的关系之下, V 的标准正交基底就是自己的固有的对偶(相互)基底. 在研究线性算子的时候, 我们要使用自然同构(将它与通常的向量空间同构$V\simeq V^{**}$做比较). 如果在V^* 上按规则

$$((\mathbf{u}|*)|(\mathbf{v}|*))^*:=(\mathbf{u}|\mathbf{v}),$$

定义纯量乘积, 同构映射Φ可以看成定理6意义下的一种度量, 所有的关于纯量乘积的公理都满足(请验证一下).

5. 标准正交基底与正交矩阵　在欧几里得向量空间V 中标准正交基底起着重要作用, 从而自然地要研究从一个标准正交基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 到另一个标准正交基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$的转换公式. 一如既往, 记

$$\mathbf{e}'_j=a_{1j}\mathbf{e}_1+a_{2j}\mathbf{e}_2+\cdots+a_{nj}\mathbf{e}_n,\quad 1\leqslant j\leqslant n, \tag{14}$$

我们就得到一个转换矩阵$A=(a_{ij})$, 在它的第k列上排列着$\mathbf{e}'_k$ 对于基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 的坐标. 到现在为止, 我们只是又把第1章§2的公式(3)再写了出来, 且在A 上有唯一的一个限制$\det A\neq 0$. 现在利用基底的标准正交性质:

$$\delta_{ij}=(\mathbf{e}'_i|\mathbf{e}'_j)=\left(\sum_k a_{ki}\mathbf{e}_k|\sum_l a_{lj}\mathbf{e}_l\right)=\sum_{k,l}a_{ki}a_{lj}(\mathbf{e}_k|\mathbf{e}_l)=\sum_k a_{ki}a_{kj}.$$

也就是

$$a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\begin{cases}0, i\neq j,\\ 1, i=j.\end{cases} \tag{15}$$

联系到转置矩阵${}^tA$, 关系式(14)(或(15))可以再写成简捷形式

$${}^tA\cdot A=E. \tag{16}$$

从而有$A^{-1}={}^tA$. 因为$A^{-1}A=E\Rightarrow A\cdot A^{-1}=E$, 所以又有

$$A\cdot{}^tA=E, \tag{16'}$$

它可以导出关系式

$$a_{i1}a_{j1}+a_{i2}a_{j2}+\cdots+a_{in}a_{jn}=\begin{cases}0, i\neq j,\\ 1, i=j.\end{cases} \tag{15'}$$

定义7 满足等价条件(15), (15′), (16), (16′)之一的矩阵$A=(a_{ij})$ 即称之为**正交矩阵**. 所有的n阶正交矩阵的集合用符号$O(n)$ 表示.

可以直接验证(而且我们还会返回到它上), $O(n)$ 是个群, 称它为**正交群**. 现在, 如果A 是任意一个正交矩阵, 那么, 由标准正交基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 按公式(14)得到的向量组$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$ 同样也是标准正交基底.

我们得到下列重要结论

定理8 由一个标准正交基底向另一个标准正交基底的转换矩阵是正交矩阵, 而且, 所有的正交矩阵都可以是这种转换矩阵.

注意, 由公式(14)和向量的标准正交性可以得到正交矩阵A 的元素a_{ij} 的几何解释. 就是

$$a_{ij}=(\mathbf{e}_i|\mathbf{e}'_j)=\cos\varphi_{ij}, \tag{17}$$

其中φ_{ij} 是原来的基底向量$\mathbf{e}_i$ 和新的基底向量$\mathbf{e}'_j$ 之间的夹角. 把等式(16)转化成行列式, 还可以得到$(\det\ A)^2=1$, 即, 所有的正交矩阵的行列式为+1或−1.

由第1章§2我们熟悉了向量$\mathbf{v}\in V$ 在基底转换时坐标的表示规则. 如果$\mathbf{x}=\sum\limits_i x_i\mathbf{e}_i=\sum\limits_i x'_i\mathbf{e}'_i$, 那么

$$x_i=\sum_j a_{ij}x'_j,\quad i=1,2,\cdots,n.$$

但是, 现在还进一步知道$A^{-1}={}^tA$, 所以

$$x'_i=\sum_j a_{ji}x_j,\quad i=1,2,\cdots,n.$$

6. 辛空间 纯量乘积的概念是“多面孔的”. 如, 在$V=\mathbb{R}^{2m}$ 上任意一个非退化的斜对称的双线性型就给出一个**辛线性结构**. 二次型本身, 常常表成$[\mathbf{x}|\mathbf{y}]$, 仍然被称为是**辛纯量乘积**. 对(V , $[*|*]$)被称为**辛空间**. 对照第1章§4的定理9, 不失一般性, 可以把辛结构认为是标准的(相对于J或J_0). 对应的基底称为**辛基底**. 我们提到过的定理9的推论有另外一种说法: 所有的维数相同的辛空间都是同构的.

如同在欧几里得空间一样, 很自然地也可以研究V 上的保持辛结构的线性算子.

定义8 称线性算子$\mathcal{A}: V\to V$ 为**辛算子**, 如果

$$[\mathcal{A}\mathbf{x}|\mathcal{A}\mathbf{y}]=[\mathbf{x}|\mathbf{y}]\quad \forall\mathbf{x},\mathbf{y}\in V=\mathbb{R}^{2m}.$$

空间V 上所有辛线性算子的集合称为**辛群**且用符号Sp$(2m)$:=Sp(V)代表. $\mathcal{A}$ 属于Sp$(2m)$的条件的矩阵形式可写成

$$ {}^tA\cdot J_0\cdot A=J_0=\begin{pmatrix}0 & -E_m\\ E_m & 0\end{pmatrix}. $$

这时, 称A 是个**辛矩阵**. 由此得出$\det\ A=\pm 1$. 事实上, $\det\ A=1$, 这可以由定理8(第1章§4的第10目)直接计算出来:

$$ 1=\mathrm{Pf}(J_0)=\mathrm{Pf}({}^tA\cdot J_0\cdot A)=(\det\ A)\mathrm{Pf}(J_0)=\det\ A. $$

直接验证就容易确信, 集合Sp$(2m)$ 实际上是个群, 也就是相对于通常的运算$(A, B)\mapsto AB, A\to A^{-1}$ 是封闭的.

当$m=1$ 时, 有同构Sp$(2)\simeq \mathrm{SL}(2,\mathbb{R})$. 实际上, 对任意$\alpha,\beta,\gamma,\delta$, 由条件

$$ {}^t\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\cdot\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\cdot\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} $$
$$ =\begin{pmatrix}0 & -(\alpha\delta-\beta\gamma)\\ \alpha\delta-\beta\gamma & 0\end{pmatrix}=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} $$

可以得到$\alpha\delta-\beta\gamma=1$.

在第5目, 我们已经看到, 正交群的元素与欧几里得向量空间的标准正交基底一一对应. 类似地, 线性算子$\mathcal{A}:\mathbb{R}^{2m}\to\mathbb{R}^{2m}$ 是辛算子, 当且仅当, 它在辛空间上导出辛基底(并且, 对任意两个辛基底都能找到一个算子, 它把一个变成另外一个).

正如我们要指出的, 辛算子(辛矩阵)的谱具有一系列有趣的特性:

定理9　*蕴涵式*

$$ \mathcal{A}\in \mathrm{Sp}(2m)\Rightarrow \chi_{\mathcal{A}}(t)=t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right) $$

成立.

特征多项式$\chi_{\mathcal{A}}(t)$ 的根可以分成四个一组(或者成对), 它们有序地以实轴对称并且以单位圆周对称.

换言之, $\chi_{\mathcal{A}}(t)=\sum\limits_{i=1}^{2m}a_it^i$ 是个递归多项式: $a_i=a_{2m-i}, i=1,2,\cdots$.

证明　设A 是算子$\mathcal{A}$ 的矩阵. 由定义的关系式${}^tAJ_0A=J_0$ 有${}^tA^{-1}=J_0AJ_0^{-1}=-J_0AJ_0$. 可见,

$$ \begin{aligned}\chi_{\mathcal{A}}(t)&=\det(tE-A)=\det(J_0(tE-A)J_0^{-1})=\det(tE-J_0AJ_0^{-1})\\ &=\det(tE-{}^tA^{-1})=\det({}^t(tE-A^{-1}))=\det(tE-A^{-1})\\ &=\det(tE-A^{-1})\det\ A=\det(tA-E)=\det(E-tA)\end{aligned} $$

$$= t^{2m}\det\left(\frac{1}{t}E - A\right) = t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right).$$

我们看到, 与辛线性算子$\mathcal{A}$ 的特征值λ 一起, λ^{-1} 也同样是个特征值. 此外, $\chi_{\mathcal{A}}(t) \in \mathbb{R}[t]$ ，与多项式$\chi_{\mathcal{A}}(t)$ 的复根λ一起, $\bar{\lambda}$ 也是它的一个根. 每个根λ ，如果$|\lambda| \neq 1$, $\mathrm{Im}\,\lambda \neq 0$, 就确定了四个根$\lambda$, $\bar{\lambda}$, $1/\lambda$, $1/\bar{\lambda}$, 得四根组(图4). 在破坏了条件其中之一时, 就得到一对根. □

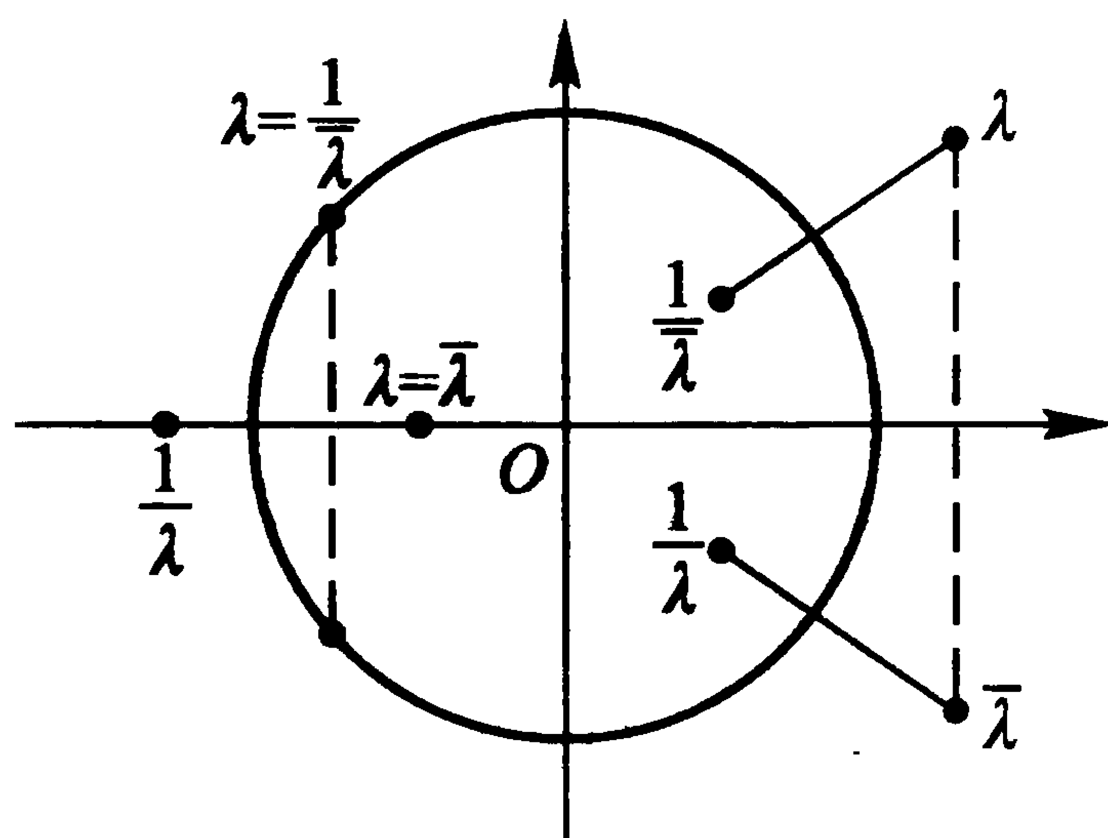


图4

有时, 对带有标准结构J_0 :

$$[\mathbf{x}|\mathbf{y}] = \sum_{i=1}^{m}(x_{i+m}y_i - x_i y_{i+m})$$

的辛空间V引入相对照的欧几里得结构:

$$(\mathbf{x}|\mathbf{y}) = \sum_{k=1}^{2m} x_k y_k$$

(尽管与J_0 的联系并不规范, 即所用的方式取决于坐标系的选择).

对于$[\mathbf{x}|\mathbf{y}]$, 如同对任意的双线性型(见§3), 可以找到一个线性算子$\mathcal{J}$, 使得

$$[\mathbf{x}|\mathbf{y}] = (\mathbf{x}|\mathcal{J}\mathbf{y}).$$

$[\mathbf{x}|\mathbf{y}]$ 的斜对称性导致算子$\mathcal{J}$ 的斜对称性. 事实上, 在选定的基底之下, 它们的矩阵就是J_0 . 显然, $\mathcal{J} \in \mathrm{Sp}(V)$ 且$\mathcal{J}^2 = -\mathcal{E}$. 直接可以看出来算子$\mathcal{J}$ 能够表示成把每个辛(双曲)平面转π /2角的旋转.

我们称一个平面$\Pi \in V$ 是**零平面**, 如果$[\Pi|\Pi] = 0$, 也就是对任意$\mathbf{x}$, $\mathbf{y} \in \Pi$, $[\mathbf{x}|\mathbf{y}] = 0$ (斜正交性). 由$\mathcal{J}$ 的定义推知, 平面Π 是个零平面, 当且仅当, 平面Π 与$\mathcal{J}(\Pi)$在欧几里得意义之下正交. 因为$\mathcal{J}$是个非退化的算子, 所以$\dim\,\Pi = \dim\,\mathcal{J}(\Pi)$, 因此, 在$V = \mathbb{R}^{2m}$ 中零平面的维数不超过m.

满足条件$\mathcal{J}^2 = -\mathcal{E}$ 的算子$\mathcal{J}$ 给出了辛空间中引入复结构的可能性. 在这个前提下的一些应该理解的东西将在§4中加以解释.

习　题

1. 在所有次数<3的多项式空间P_3中, 对于由公式$(f|g)=\int_{-1}^{1}f(t)g(t)dt$给定的纯量乘积(公式(2)的特殊情形), 向量1和向量t是正交的. 找出:

(1) 子空间$\langle 1,t\rangle^{\perp}$;

(2) P_3的一个标准正交基底.

2. 设$(V,(*|*))$是3维欧几里得向量空间, 对任意$\mathbf{x}\in V$, $\|\mathbf{x}\|^2=(\mathbf{x}|\mathbf{x})=3x_1^2+2x_2^2+x_3^2-4x_1x_2-2x_1x_3+2x_2x_3$ (验证, 这个二次型是正定的), 试找出:

(1) 向量$\mathbf{x}=[1,1,1]$, $\mathbf{y}=[2,2,1]$之间的夹角α;

(2) 所有与$\mathbf{x}$正交的向量.

3. 运用格拉姆–施密特正交化过程证明, 任意一个非退化矩阵$A=(a_{ij})\in M_n(\mathbb{R})$都可以写成乘积$A=BC$的形式, B是正交矩阵, C是上三角形矩阵, 且$\pm\det\ A=\det\ C$.

4. 把集合$M_n(\mathbb{R})$看成是带有标准纯量乘积的欧几里得空间, 维数为n^2. 它包含一个正交矩阵群$O(n)$, 每个矩阵决定了$n(n+1)/2$个关系式(15)或(15′) ($O(n)$就是正交群). 这样一来, 可以把$O(n)$当成一个$n^2-n(n+1)/2=n(n-1)/2=\dim\mathfrak{o}\ (n)$维的“代数流形”, 其中$\mathfrak{o}(n)$是$n$阶斜对称矩阵的向量空间. 在$O(n)$和$\mathfrak{o}(n)$之间自然地产生了一个相当好的对应关系, 凯莱变换

$$K=(E-A)^{-1}(E+A),\quad A=(E-K)^{-1}(E+K),\tag{18}$$

就是这种对应的一个例子.

需要证明, 如果$A\in O(n)$, $\det(E-A)\neq 0$, 那么$K\in\mathfrak{o}(n)$. 反之, 验证, 每个矩阵$K\in\mathfrak{o}(n)$都对应一个$1\notin\operatorname{Spec}(A)$的矩阵$A$. 在上述提出的“代数流形”$O(n)$的对应中需要去掉方程$\det(E-A)=0$决定的“超平面”.

可以提出另一个凯莱变换

$$K=(E+A)^{-1}(E-A),\quad A=(E+K)^{-1}(E-K),\tag{19}$$

其中需要从$O(n)$中排除满足方程$\det(A+E)=0$的“超平面”, 或者, 等价于$-1\notin\operatorname{Spec}(A)$的正交矩阵.

要求证实关系(19)的正确性.

5. 验证, 由斜对称矩阵用凯莱变换(在关系式(18)和(19)中)得到的正交矩阵的行列式均为1(用符号$SO(n)$代表所有这样的矩阵的集合).

6. 证明, $n\times n$阶正交矩阵A的特征多项式$\chi_A(t)$具有性质:

$$t^n\chi_A(1/t)=\pm\chi_A(t).$$

7. 设$A=\left[A_{(1)},\cdots,A_{(n)}\right]$是一个由相互正交的行排成的矩阵, 证明

$$|\det\ A|=\left\|A_{(1)}\right\|\cdot\left\|A_{(2)}\right\|\cdots\left\|A_{(n)}\right\|$$

(向量在$\mathbb{R}^n$的标准模).

8. 设$X=\left[X_{(1)},\cdots,X_{(n)}\right]$是$M_n(\mathbb{R})$中的任意一个矩阵, 证明

$$|\det X|\leqslant\left\|X_{(1)}\right\|\cdot\left\|X_{(2)}\right\|\cdots\left\|X_{(n)}\right\|$$

(阿达马不等式).

§2 埃尔米特向量空间

1. 埃尔米特型 很多问题与作用在复向量空间上的线性算子有这样或那样的联系. 由于这个原因, 复向量空间赢得人们的特别关注. $\mathbb{R}$ 上欧几里得空间上各种度量关系自然地充当了在复的情形引入纯量乘积的刺激因素. 但是, 正如在第1章§4的第7目的末尾指出的那样, 标准双线性型$s(\mathbf{x},\mathbf{y}) = x_1y_1 + \cdots + x_ny_n,\ x_j, y_j \in \mathbb{C}$ 却不能是达到此目的的一个出发点, 因为向量$\mathbf{x} \in \mathbb{C}^n$ 的长度(模)$\|\mathbf{x}\| = \sqrt{s(\mathbf{x},\mathbf{x})}$ 具有"令人厌恶的"性质:

$$\|i\mathbf{x}\|^2 = s(i\mathbf{x}, i\mathbf{x}) = i^2 s(\mathbf{x},\mathbf{x}) = -\|\mathbf{x}\|^2.$$

如果$\mathbf{x} \neq \mathbf{0}$ 且$\|\mathbf{x}\| > 0$, 那么$i\mathbf{x} \in \mathbb{C}^n$, 但$\|i\mathbf{x}\| < 0$.

要是我们想利用直观理解的向量长度的概念, 上面的定义显然是不能接受的.

埃尔米特空间(酉空间)可以充当欧几里得向量空间的极好的类似对象. 我们引入下面的.

定义1 称$f : V \times V \to \mathbb{C}$ 是复向量空间上的一个**半双线性型**, 如果

i) $f(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha f(\mathbf{x},\mathbf{z}) + \beta f(\mathbf{y},\mathbf{z}),\ \forall \alpha,\beta \in \mathbb{C}$, $\mathbf{x},\mathbf{y},\mathbf{z} \in V$, 也就是说, 在第二个变量固定时, f 对第一个变量是线性的;

ii) $f(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \overline{\alpha} f(\mathbf{x},\mathbf{y}) + \overline{\beta} f(\mathbf{x},\mathbf{z})$, 其中α,β 上的小横杠表示通常的共轭复数(在第一个变量固定时, 对第二个变量有半线性性).

称半双线性型f为**埃尔米特型**, 如果

$$f(\mathbf{y},\mathbf{x}) = \overline{f(\mathbf{x},\mathbf{y})}. \tag{1}$$

设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 是空间V 的一个基底, 如果$\mathbf{x} = \sum\limits_i x_i\mathbf{e}_i$, $\mathbf{y} = \sum\limits_j y_j\mathbf{e}_j$, 那么

$$f(\mathbf{x},\mathbf{y}) = \sum_{i,j} f_{ij} x_i \bar{y}_j, \quad f_{ij} = f(\mathbf{e}_i,\mathbf{e}_j)$$

是型f 在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 之下值的表达式. 型f 的埃尔米特性质意味着它的矩阵$F = (f_{ij})$ 的系数满足条件

$$f_{ij} = \overline{f}_{ji}.$$

换言之.

$$F^* = F, \tag{1'}$$

其中$F^* := \overline{{}^tF}$. 满足条件(1′)的矩阵F同样也称为是**埃尔米特的**.

如果F'是埃尔米特型f在基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$之下的矩阵, 借助于转换矩阵$A$, 在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下又可以得到它, 于是有

$$F' = {}^tA \cdot F \cdot \overline{A} \tag{2}$$

($\overline{A} = (\overline{a}_{ij})$; 可以和第1章§4的表达式(5)相比较). 利用(1′)可以直接验证

$$(F')^* = {}^t\left({}^t\overline{A} \cdot \overline{F} \cdot A\right) = {}^tA \cdot {}^t\overline{F} \cdot \overline{A} = {}^tA \cdot F^*\overline{A} = {}^tA \cdot F \cdot \overline{A} = F',$$

也就是说, 这正是我们所期待的, 在用${}^tA\cdot F\cdot\overline{A}$替换$F$时仍保持埃尔米特性质.

埃尔米特型$f(\mathbf{x},\mathbf{y})$自然地对应埃尔米特二次型$f(\mathbf{x},\mathbf{x})$, 因为

$$f(\mathbf{x},\mathbf{x})=\overline{f(\mathbf{x},\mathbf{x})},$$

所以埃尔米特二次型只取实数值.此时, 如果$f(\mathbf{x},\mathbf{x})\geqslant 0$而且$f(\mathbf{x},\mathbf{x})=0\Rightarrow\mathbf{x}=\mathbf{0}$, 那么, 就说型$f$是**正定型**. 把$f$写成

$$f(\mathbf{x},\mathbf{y})=g(\mathbf{x},\mathbf{y})+ih(\mathbf{x},\mathbf{y}),$$

其中g, h都是实值函数, 利用(1), 我们容易相信, g和h都是V上的双线性型, 而且g是对称型, 而h是斜对称型. 最后, f的正定性等价于g的正定型.

定义2 域$\mathbb{C}$上一个有限维向量空间V附加了一个正定的埃尔米特型$(\mathbf{x}|\mathbf{y}):=f(\mathbf{x},\mathbf{y})$, 则称为**埃尔米特空间**(**酉空间**). 复数$(\mathbf{x}|\mathbf{y})$称为是向量$\mathbf{x},\mathbf{y}\in V$的**纯量积**(也说是**内积**).

换言之, 在新的设定之下, 我们有

$$(\mathbf{x}|\mathbf{y})=\overline{(\mathbf{y}|\mathbf{x})},$$
$$(\alpha\mathbf{x}+\beta\mathbf{y}|\mathbf{z})=\alpha(\mathbf{x}|\mathbf{z})+\beta(\mathbf{y}|\mathbf{z}),$$
$$(\mathbf{x}|\mathbf{x})\geqslant 0;\ (\mathbf{x}|\mathbf{x})=0, \text{只有}\mathbf{x}=\mathbf{0}\text{才行}.$$

例1 设

$$(\mathbf{x}|\mathbf{y})=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n}, \tag{3}$$

毫无疑义地, 我们得到了一个带有单位矩阵$F=E$的正定的埃尔米特型, 也就是复的坐标空间$\mathbb{C}^n$, 附加了这个埃尔米特型就成为埃尔米特空间了. 如果借助于矩阵A把标准正交基底变成另外一个基底, 那么, 按照(2), 就可以用我们的标准的埃尔米特型对应埃尔米特矩阵$F'={}^tA\cdot\overline{A}$.

在实的情形, 复共轭性可以忽略不计, 因此, 埃尔米特空间就是地地道道的欧几里得空间了, 如同在欧几里得空间一样, 用等式

$$\|\mathbf{v}\|=\sqrt{(\mathbf{v}|\mathbf{v})}$$

定义向量$\mathbf{v}\in V$的长度$\|\mathbf{v}\|$.

2. 度量关系 容易验证的关系式

$$2(\mathbf{u}|\mathbf{v})=\|\mathbf{u}+\mathbf{v}\|^2+i\|\mathbf{u}+i\mathbf{v}\|^2-(1+i)\left\{\|\mathbf{u}\|^2+\|\mathbf{v}\|^2\right\}$$

表明, 纯量乘积可以直接用长度术语表达出来(配极化过程). 由显然的等式

$$\|\lambda\mathbf{x}\|=\sqrt{(\lambda\mathbf{x}|\lambda\mathbf{x})}=\sqrt{|\lambda|^2(\mathbf{x}|\mathbf{x})}=|\lambda|\sqrt{(\mathbf{x}|\mathbf{x})}$$

可以推出在欧几里得空间情形已知的关于模的性质

$$\|\lambda \mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|. \tag{3'}$$

它可以被平行地推广成一些更宽泛的命题.特别地，柯西–布尼亚科夫斯基不等式(也称它是**施瓦茨不等式**)具有如下形式

$$|(\mathbf{x}|\mathbf{y})| \leqslant \|\mathbf{x}\| \cdot \|\mathbf{y}\| \tag{4}$$

(等式只有当$\mathbf{x}, \mathbf{y}$成比例时才能达到).

证明 实际上，把复数记为三角形式$(\mathbf{x}|\mathbf{y}) = |(\mathbf{x}|\mathbf{y})|\, e^{i\varphi}$, $\varphi \in \mathbb{R}$.我们可以看出，对任意$t \in \mathbb{R}$, 不等式

$$\|\mathbf{x}\|^2 t^2 + \left((\mathbf{x}|\mathbf{y})\, t^{-i\varphi} + \overline{(\mathbf{x}|\mathbf{y})} e^{i\varphi}\right) t + \|\mathbf{y}\|^2 = \left(\mathbf{x}t + \mathbf{y}e^{i\varphi} | \mathbf{x}t + \mathbf{y}e^{i\varphi}\right) \geqslant 0$$

成立. 因为$(\mathbf{x}|\mathbf{y})\, e^{-i\varphi} = |(\mathbf{x}|\mathbf{y})| = \overline{(\mathbf{x}|\mathbf{y})}\, e^{i\varphi}$, 所以，又可以把它写成形如

$$\|\mathbf{x}\|^2 t^2 + 2\,|(\mathbf{x}|\mathbf{y})|\, t + \|\mathbf{y}\|^2 \geqslant 0.$$

从判别式上得到的条件就导出所要的不等式. 当有适当的$t_0 \in \mathbb{R}$使$\mathbf{x}t_0 + \mathbf{y}e^{i\varphi} = 0$时，即$\mathbf{x}, \mathbf{y}$成比例时，它就变成了严格的等式. □

由不等式(4)可直接导出三角不等式

$$\|\mathbf{x} \pm \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \tag{5}$$

以及它的一个显然的推广

$$\|\mathbf{x} - \mathbf{z}\| \leqslant \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|.$$

例2 在$\mathbb{C}$上向量空间$C_2(a, b)$和P_n上附加纯量乘积

$$(f|g) = \int_a^b f(t)\,\overline{g(t)}\, dt,$$

显然是个埃尔米特空间.在这种情况下，不等式(5)就变成

$$\sqrt{\int_a^b |f(t) \pm g(t)|^2\, dt} \leqslant \sqrt{\int_a^b |f(t)|^2\, dt} + \sqrt{\int_a^b |g(t)|^2\, dt}$$

的形式(与§1例子中的闵可夫斯基不等式比较).在带有标准纯量乘积(3)的埃尔米特空间$\mathbb{C}^n$中不等式

$$\sqrt{\sum_{i=1}^n |x_i \pm y_i|^2} \leqslant \sqrt{\sum_{i=1}^n |x_i|^2} + \sqrt{\sum_{i=1}^n |y_i|^2}$$

成立.

运用不等式(4)可以证实, 有唯一的一个角$\varphi, 0 \leqslant \varphi \leqslant \frac{\pi}{2}$使得

$$\cos\varphi = \frac{|(\mathbf{x}|\mathbf{y})|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}.$$

量$\cos^2\varphi$的量子力学解释可在教学参考书 [2] 中找到.

3. 正交性 与实的情形一样, 在埃尔米特空间$(V, (*|*))$选出向量$\mathbf{e}_1, \cdots, \mathbf{e}_m$, 如果$(\mathbf{e}_i|\mathbf{e}_j) = \delta_{ij}$, 就称它们是**标准正交**的. 这个向量组是线性无关的, 而且可以扩充成V的标准正交基底. 为了确信这一点, 把每个向量$\mathbf{u}$与向量$\mathbf{v} = \mathbf{u} - \sum_{i=1}^{m} (\mathbf{u}|\mathbf{e}_i)\,\mathbf{e}_i$一起讨论, 而且注意到$\mathbf{v} \in \langle \mathbf{e}_1, \cdots, \mathbf{e}_m \rangle^{\perp}$, 可以把格拉姆–施密特正交化过程(见§1的第3目). 向量$\mathbf{v}$可被标准化而继续进一步的过程.顺便说一句, 对每个子空间$W \subset V$都有

$$V = W \oplus W^{\perp}, \quad W^{\perp\perp} = W. \tag{6}$$

提议将下列命题的证明作为一个不大的练习题.

定理1 设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是埃尔米特向量空间(或欧几里得向量空间)$(V, (*|*))$的一个标准正交基底.

那么:

i) 对所有的$\mathbf{x} \in V$, $\mathbf{x} = \sum_i (\mathbf{x}|\mathbf{e}_i)\mathbf{e}_i$;

ii) 对任意$\mathbf{x}, \mathbf{y} \in V$, $(\mathbf{x}|\mathbf{y}) = \sum_i (\mathbf{x}|\mathbf{e}_i)\,(\mathbf{e}_j|\mathbf{y})$(帕塞瓦尔等式);

iii) $\mathbf{x} \in V \Rightarrow \|\mathbf{x}\|^2 = \sum_i |(\mathbf{x}|\mathbf{e}_i)|^2$.

设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是埃尔米特空间 V的一个标准正交基底, 在定理1中使用了下面的表达式. 对任意$\mathbf{x} = \sum_i x_i\mathbf{e}_i$, 借助于纯量乘积对第一个变量的线性性质, 有

$$(\mathbf{x}|\mathbf{e}_j) = \left(\sum_i x_i\mathbf{e}_i \,\middle|\, \mathbf{e}_j\right) = \sum_i x_i\,(\mathbf{e}_i|\mathbf{e}_j) = x_j.$$

这样一来, 我们导出了线性函数$f_j = (*|\mathbf{e}_j) : V \to \mathbb{C}$, 它把每个向量$\mathbf{x} = \sum_i x_i\mathbf{e}_i$与它对于基底$(\mathbf{e}_i)$的第$j$个坐标相提并论. 现在, 如果$\mathbf{y} = \sum_j y_j\mathbf{e}_j$也是$V$的一个向量. 那么,

$$(\mathbf{x}|\mathbf{y}) = \sum_{i,j} x_i\overline{y_j}\,(\mathbf{e}_i|\mathbf{e}_j) = x_1\overline{y_1} + \cdots + x_n\overline{y_n},$$

这也就是埃尔米特空间V按着对$\mathbb{C}^n$的标准纯量乘积的公式(3)选定标准正交基底之下的纯量乘积的算法. 用这种思想可以定义埃尔米特空间的同构 $\mathbb{C}^n \simeq V : (x_1, \cdots, x_n) \to \sum_i x_i\mathbf{e}_i$, 它是个双射, 而且保持纯量乘积. 与欧几里得向量空间不同, 不能把

埃尔米特空间和它的对偶空间等同起来, 需要把按下面定义的意义下的半线性函数与线性函数统一加以研究.

定义3 设f是复向量空间V上的一个通常的线性型(函数).称满足条件

$$\overline{f}(\mathbf{x}+\mathbf{y})=\overline{f}(\mathbf{x})+\overline{f}(\mathbf{y}),\quad \overline{f}(\lambda\mathbf{x})=\overline{\lambda}\,\overline{f}(\mathbf{x})$$

的函数$\overline{f}:V-\mathbb{C}$是与$f$**共轭的线性函数**(或**半线性函数**).

如果$(V,(*|*))$是个埃尔米特空间, 那么f可以写成$f(\mathbf{x})=(\mathbf{x}|\mathbf{a})$的形式(与§1的定理9比较), 但$f$与$\mathbf{a}$的对应并非是线性的. 现在, 如果$\overline{f}$是个半线性函数, 那么, 在$V$中选择某个标准正交基底$(\mathbf{e}_i)$且令$\mathbf{a}=\sum\limits_i\overline{f}(\mathbf{e}_i)\,\mathbf{e}_i$, 我们就有对任意$\mathbf{x}=\sum\limits_j x_j\mathbf{e}_j$的关系式

$$(\mathbf{a}|\mathbf{x})=\sum_i\left(\mathbf{e}_i|\sum_j x_j\mathbf{e}_j\right)=\sum_i\overline{f}(\mathbf{e}_i)\,\overline{x_i}=\overline{f}(\mathbf{x}).$$

显然, 由型$(*|*)$的正定性可导出向量$\mathbf{a}$的唯一性. 由于$(*|*)$的埃尔米特性质, 允许记

$$\overline{f}(\mathbf{x})=(\mathbf{a}|\mathbf{x})=\overline{(\mathbf{x}|\mathbf{a})}=\overline{f(\mathbf{x})}.$$

4. 酉矩阵 在欧几里得向量空间中借助正交矩阵可实现从一个标准正交基底向另外一个标准正交基底的转化(§1定理10). 在埃尔米特空间情形有类似的命题. 设$(\mathbf{e}_i)$和$(\mathbf{e}'_j)$是埃尔米特空间$(V,(*|*))$的两个标准正交基底, 它们以转换矩阵$A=(a_{ij})$相联系: $\mathbf{e}'_j=\sum\limits_i a_{ij}\mathbf{e}_i$. 那么,

$$\delta_{jk}=\left(\mathbf{e}'_j|\mathbf{e}'_k\right)=\sum_{i,s}a_{ij}\overline{a_{sk}}\,(\mathbf{e}_i|\mathbf{e}_s)=\sum_i a_{ij}\overline{a_{ik}}.$$

换言之,

$$A\cdot A^*=E=A^*\cdot A, \tag{7}$$

其中$A^*:={}^t\overline{A}$是与A**共轭的埃尔米特矩阵**(回顾, $\overline{A}=(\overline{a_{ij}})$).

定义4 满足条件(7)的矩阵A被称为是**酉矩阵**.

显然, 在实的情形, 酉矩阵就是正交矩阵. 其次, $\det\,\overline{A}=\overline{\det\,A}$, 所以, $\det\,A^*=\overline{\det\,A}$, 再注意到(7), 我们得到$|\det\,A|=1$, **也就是说, 对任意酉矩阵**$A$, 有$\det\,A=e^{i\varphi}$. 特别地, 酉矩阵是非退化的.

由A^*的定义直接得到

$$(A\cdot B)^*=B^*\cdot A^*. \tag{8}$$

所以, 在A, B都是酉矩阵的情况下, 我们可以推导出关于它们乘积的复正交性质: $AB\,(AB)^*=A\,(B\,B^*)\,A^*=AEA^*=AA^*=E$.其次, $A^{-1}A=E=AA^{-1}\Rightarrow A^*\left(A^{-1}\right)^*=E=\left(A^{-1}\right)^*A^*\Rightarrow A^{-1}\left(A^{-1}\right)^*=E=\left(A^{-1}\right)^*A^{-1}$, 也就是说, A^{-1}与A一样也是酉矩阵.当然, 这与对正交矩阵的了解是一样的.

注意到群的一般定义, 我们可以看出, 有

定理2 下列命题真确:

i) 所有的n阶酉矩阵都是(酉)群$U(n)$的元素;

ii) 酉群$U(n)$包含由所有n阶实的正交矩阵组成的正交群$O(n)$作为它的一个子群;

iii) 行列式为1的正交(相应地, 复正交)矩阵构成特殊正交子群$SO(n)$(相应地, 特殊酉群$SU(n)$).

这样一来,

$$SO(n) = O(n) \cap SL(n) \subset SU(n) = U(n) \cap SL(\mathbb{C}).$$

一般地说, 在任意域$\mathfrak{K}$上定义正交群$O(n,\mathfrak{K})$可能没有特别的困难, 与此相同, 补充了类似于复共轭$\alpha \mapsto \overline{\alpha}$, 也可定义任意域$\mathfrak{K}$上的酉群$U(n,\mathfrak{K})$.

5. 可赋范的向量空间 §1中的不等式(10)(三角形的一个边的长度不超过它的另外两个边的长度之和)以及它在酉空间的类似的不等式(5)允许把带有纯量乘积的向量空间在下面定义的意义下看成是一个可度量的空间.

定义5 设E是某些点的集合, 而$d : E \times E \to \mathbb{R}$是个映射, 同时, 对任意两个点$u$, $v \in E$, 有非负实数$d(u,v)$ (u, v之间的距离)和它们相对应, 而且还满足下列性质:

i) $d(u,v) = d(v,u)$(对称性);

ii) $d(u,v) = 0 \Leftrightarrow u = v$;

iii) $d(u,w) \leqslant d(u,v) + d(v,w)$(三角不等式).

具有这种性质的函数d可称之为**度量**, 称对(E,d)是个**度量空间**.

例3 在带有纯量乘积的向量空间V上, 进一步以其本身确定的模$\|\mathbf{x}\|$取$d(\mathbf{x},\mathbf{y}) := \|\mathbf{x}-\mathbf{y}\|$作为向量$\mathbf{x},\mathbf{y}$之间的距离.比如, 对于$V = C_2(a,b)$, 可用

$$d(f,g) = \sqrt{\int_a^b |f(t) - g(t)|^2 \, dt}$$

充当度量.而同样地, 函数

$$d'(f,g) = \max_{a \leqslant t \leqslant b} |f(t) - g(t)|, \quad d''(f,g) = \int_a^b |f(t) - g(t)| \, dt,$$

也满足定义5中的条件i)— iii), 直接验证就可以确信.

度量出来的量立刻就可以导出拓扑学与分析学中最简单的概念, 包括极限过程(取极限).度量空间$(E,\ d)$的子空间

$$B(a_0,r) = \{x \in E | d(a_0,x) < r\},$$
$$\overline{B}(a_0,r) = \{x \in E | d(a_0,x) \leqslant r\},$$
$$S(a_0,r) = \{x \in E | d(a_0,x) = r\}$$

分别称为是以点a_0为中心, 以r为半径的**开球**, **闭球**和**球面**.

称子集$F \subset E$是**有限制的**, 如果它被包含在一个半径$r < \infty$的球中.

称(E,d)中点的序列$e_1, e_2, \cdots, e_n, \cdots$**收敛到**点$e \in E$, 如果$\lim\limits_{n\to\infty} d(e_n, e) = 0$.称序列是**基本的**或**柯西序列**, 如果对任意$\varepsilon > 0$, 均有$N = N(\varepsilon)$使得当$m, n > N$时必有$d(e_n, e_m) < \varepsilon$.称度量空间$E$是**完备的**, 如果它里边的每个柯西序列都是收敛的.由在分析学中已经证明了的$\mathbb{R}$和$\mathbb{C}$的完备性可以得出, 带有三个度量

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{\frac{1}{2}},$$

$$d_1(\mathbf{x}, \mathbf{y}) = \max(|x_i - y_i|), \quad d_2(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

中任意一个度量的空间$\mathbb{R}^n$和$\mathbb{C}^n$都是完备的(验证, d_1和d_2都是度量, 而对于d, 可由例2得出).

也就是说, 设V是个实的或者复的带有度量d的向量空间, 在d满足下面两个补充条件时, 特别重要:

iv) 对任意$\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}+\mathbf{z}, \mathbf{y}+\mathbf{z})$(相对于平移的不变性);

v) $d(\lambda\mathbf{x}, \lambda\mathbf{y}) = |\lambda|\, d(\mathbf{x}, \mathbf{y})$(乘以纯量$\lambda$等于将距离延长$|\lambda|$倍).

定义6 用$\|\mathbf{x}\|$代表数$d(\mathbf{x}, \mathbf{0})$, 并把它称为$\mathbf{x} \in V$对于具有性质iv), v)的度量$d$的**模**.

我们已经用特殊的方式(例3)在带有纯量乘积$(*|*)$的空间中引进了度量d, 而且向量$\mathbf{x}$的新模和原来的模是一致的. 所以, 我们仍然使用从前的符号$\|\mathbf{x}\|$. 回到一般情形, 我们应当证明, 模满足下列性质:

$$\|\mathbf{0}\| = 0; \text{如果}\mathbf{x} \neq \mathbf{0}; \text{ 那么 } \|\mathbf{x}\| > 0;$$

$$\|\lambda\mathbf{x}\| = |\lambda|\, \|\mathbf{x}\|; \text{ 对所有的}\lambda \in \mathbb{C}, \mathbf{x} \in V;$$

$\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$; 对所有的$\mathbf{x}, \mathbf{y} \in V$.

前两个性质可由度量公理和条件iv), v)直接推导出来; 第三个性质可以这样验证: $\|\mathbf{x}+\mathbf{y}\| = d(\mathbf{x}+\mathbf{y}, \mathbf{0}) = d(\mathbf{x}, -\mathbf{y}) \leqslant d(\mathbf{x}, \mathbf{0}) + d(\mathbf{0}, -\mathbf{y}) = \|\mathbf{x}\| + \|\mathbf{y}\|$.

定义7 向量空间V, 附加了满足上述三个条件的模函数$\|*\| : V \to \mathbb{R}$, 即称为**赋范空间**. 完备的赋范向量空间被称为**巴拿赫空间**.

空间$\mathbb{R}^n$和$\mathbb{C}^n$分别带有上面见到的任意模都是一个巴拿赫空间. 还要注意, 按模来恢复度量: 设$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x}-\mathbf{y}\|$, 很容易验证度量公理, 对此, 有$d(\mathbf{x}, \mathbf{0}) = \|\mathbf{x}\|$.

前面我们给出的度量空间中序列的收敛性概念专门用到可赋范的向量空间情形, 称之为**按模收敛**.

容易证实

定理3 设V是$\mathbb{R}$或者$\mathbb{C}$上的带有纯量乘积的n维向量空间.

那么, 向量序列$\mathbf{x}_k \in V$, $k=1,2,\cdots$收敛到$\mathbf{x}\in V$的下述两个收敛性概念是等价的:

i) 当$k\to\infty$时, $\|\mathbf{x}_k-\mathbf{x}\|\to 0$;

ii) 当$k\to\infty$时, 对任意固定的$\mathbf{y}\in V$, $(\mathbf{x}_k-\mathbf{x}|\mathbf{y})\to 0$.

证明　i)⇒ ii). 因为借助不等式(4)有

$$|(\mathbf{x}_k-\mathbf{x}|\mathbf{y})|\leqslant\|\mathbf{x}_k-\mathbf{x}\|\cdot\|\mathbf{y}\|\to 0.$$

ii)⇒ i). 为确信这一点, 在V中取一标准正交基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$. 如果ii)真确, 那么, 对每个$i=1,2,\cdots,n$都有$(\mathbf{x}_k-\mathbf{x}|\mathbf{e}_i)\to 0$. 因此, 用等式

$$\|\mathbf{x}_k-\mathbf{x}\|^2=\sum_{i=1}^{n}|(\mathbf{x}_k-\mathbf{x}|\mathbf{e}_i)|^2$$

(定理1, iii)), 即可归于结论, $\|\mathbf{x}_k-\mathbf{x}\|\to 0$. □

线性结构允许定义一个比用部分和的模定义收敛性更强的收敛性概念.即, 称级数$\sum\limits_{i=1}^{\infty}\|\mathbf{x}_i\|$是**绝对收敛的**, 如果级数$\sum\limits_{i=1}^{\infty}\|\mathbf{x}_i\|$是收敛的.

习　题

1. $\mathbb{C}O(n)=\{A\in M_n(\mathbb{C})\,|\,{}^tA\cdot A=E\}$是复正交群的定义. 显然, $O(n)\subset\mathbb{C}O(n)$且$SO(n)\subset S\mathbb{C}O(n)$(后二者分别是$O(n)$和$\mathbb{C}O(n)$中行列式为1的矩阵的子群). 能够仿照定理2, 说有结论$\mathbb{C}O(n)\subset U(n)$和$S\,\mathbb{C}O(n)\subset U(n)$吗?

2. 给出用$\mathbb{R}^n$上任意纯量积派生出来的第5目中的度量d_1和d_2(在$\mathbb{C}$中亦类似).

3. 利用基本工具验证, 在空间$\mathbb{R}^n$上对任意$p\geqslant 1$公式

$$\|\mathbf{x}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}$$

给出了一个模(称为L_p–模). 再验证

$$\lim_{p\to\infty}\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}=\max_{1\leqslant i\leqslant n}|x_i|.$$

于是, 可以形式地认为, $d_1(\mathbf{x},\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|_\infty$(显然, $d_2(\mathbf{x},\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|_1$, $d(\mathbf{x},\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|_2$).

在向量空间$C(0,1)$中连续函数f: $[0,1]\to\mathbb{R}$有类似的模

$$\|f\|_\infty=\max_{0\leqslant t\leqslant 1}|f(t)|,\quad \|f\|_1=\int_0^1|f(t)|dt.$$

更一般地: $\|f\|_p=\left(\int_0^1|f(t)|^p\,dt\right)^{1/p}$(这个习题对特征无要求).

§3 带有纯量乘积的空间上的线性算子

1. 线性算子与θ线性型之间的关系 所谓向量空间V上的θ线性型，当V是实的向量空间时，即理解为双线性型(θ=2)，当V是复的向量空间时，即理解为半双线性型$\left(\theta=\dfrac{3}{2}\right)$. 我们现在假定$V$是$\mathbb{R}$上带有纯量乘积$(*|*)$的欧几里得空间(相应地，在$\mathbb{C}$上就是埃尔米特空间). 其次，设$\mathcal{A}$是$V$上的任意一个线性算子. 在第2章§3的第6目已经引进了V^*上与$\mathcal{A}$共轭的线性算子$\mathcal{A}^*$的概念. 在带有纯量乘积的空间的情形，在线性算子和θ线性型之间有更进一步的类似物，反映出在V和V^*之间自然映射的存在(至少在实的情形如此)，并且也能反映出$\mathcal{A}^*$在V上的直接影响. 说到这里先停一下则更为方便. 研究映射

$$f_{\mathcal{A}}: V\times V\to\mathfrak{K}\ \left(\mathfrak{K}=\mathbb{R}\text{或者}\mathfrak{K}=\mathbb{C}\right),$$

它的定义规则是

$$f_{\mathcal{A}}(\mathbf{x},\mathbf{y}):=(\mathcal{A}\mathbf{x}|\mathbf{y}). \tag{1}$$

由纯量乘积的性质可以直接推出，$f_{\mathcal{A}}$是V上的θ线性型，即，在实的情形它是个双线性型而在复的情形是个半双线性型.类似的检查我们将不再重复.

显然，对应$\mathcal{A}\mapsto f_{\mathcal{A}}$给出了$\mathcal{L}(V)\to\mathcal{L}_\theta(V,\mathfrak{K})$的一个内射. 事实上，如果$(\mathcal{A}\mathbf{x}|\mathbf{y})=f_{\mathcal{A}}(\mathbf{x},\mathbf{y})=(\mathcal{B}\mathbf{x}|\mathbf{y})$，那么，$((\mathcal{A}-\mathcal{B})\mathbf{x}|\mathbf{y})=(\mathcal{A}\mathbf{x}-\mathcal{B}\mathbf{x}|\mathbf{y})=(\mathcal{A}\mathbf{x}|\mathbf{y})-(\mathcal{B}\mathbf{x}|\mathbf{y})=0$，$\forall\mathbf{y}\in V$，从而$(\mathcal{A}-\mathcal{B})\mathbf{x}=\mathbf{0}$，$\forall\mathbf{x}\in V$，或者，$\mathcal{A}=\mathcal{B}$. 进而，再由等式$\dim\ \mathcal{L}(V)=\dim\ \mathcal{L}_\theta(V,\mathfrak{K})$ 得出我们定义的这个映射的双射性.

话又说回来了，这一点也可以按照给定的θ线性型$f(\mathbf{x},\mathbf{y})$得到线性算子$\mathcal{A}_f$的显式构作看出来:

$$f(\mathbf{x},\mathbf{y})=(\mathcal{A}_f\mathbf{x}|\mathbf{y}), \tag{2}$$

需要做到的就是这些. 设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是$V$的一个标准正交基底，而$F$是$\theta$线性型$f(\mathbf{x},\mathbf{y})$在这个基底之下的对应的矩阵. 一如既往，用$X=[x_1,\cdots,x_n]$表示向量$\mathbf{x}=\sum\limits_i x_i\mathbf{e}_i$的坐标列，鉴于基底的标准正交性质，向量$\mathbf{x}$和向量$\mathbf{y}=\sum\limits_j y_j\mathbf{e}_j$的纯量乘积可以表示成行${}^tX$和列$\overline{Y}$的乘积形式: $(\mathbf{x}|\mathbf{y})={}^tX\cdot\overline{Y}$.

我们取$\mathcal{A}_f$作为以${}^tF$为矩阵的线性算子. 它在基底$(\mathbf{e}_i)$之下的坐标列的变换$X\to{}^tFX$. 现在，定义的关系式(2)就是刚刚引进的表示的一个简单的解释:

$$f(\mathbf{x},\mathbf{y})={}^tXF\overline{Y}={}^t\left({}^tFX\right)\overline{Y}=(\mathcal{A}_f\mathbf{x}|\mathbf{y}).$$

这样，我们就可以将某个线性算子$\mathcal{A}_f^*$与矩阵$\overline{F}$联系到一起. 于是，

$$f(\mathbf{x},\mathbf{y})={}^tX\left(\overline{\overline{F}Y}\right)=\left(\mathbf{x}|\mathcal{A}_f^*\mathbf{y}\right).$$

如果设$A = {}^tF$, $A^* = {}^t\overline{A} = \overline{F}$, 那么, A就是我们的线性算子$\mathcal{A}_f$的矩阵, 而$A^* = {}^t\overline{A}$就是算子$\mathcal{A}_f^*$的矩阵.

所有的这些叙述证实了下面的论断:

定理1 设V是个带有纯量乘积$(*|*)$的向量空间. 那么, 型

$$f_{\mathcal{A}}(\mathbf{x},\mathbf{y}) = (\mathcal{A}\mathbf{x}|\mathbf{y}), \quad f_{\mathcal{A}}(\mathbf{x},\mathbf{y}) = (\mathbf{x}|\mathcal{A}^*\mathbf{y}) \tag{3}$$

中的任意一个都建立了V上θ线性型与V上算子空间之间的一个双射对应.(3)中的两个公式一起唯一地确定了与$\mathcal{A}$共轭的线性算子$\mathcal{A}^*: V \to V$.

在标准正交基底之下, 算子$\mathcal{A}^*$的矩阵可以通过将$\mathcal{A}$的矩阵转置并取共轭复数(在$\mathfrak{K} = \mathbb{C}$的情形)得到.

当$\mathfrak{K} = \mathbb{R}$时, 定义

$$(\mathcal{A}\mathbf{x}|\mathbf{y}) = (\mathbf{x}|\mathcal{A}^*\mathbf{y}), \tag{4}$$

就和第2章§3的第6目共轭算子的定义完全一致了, 因为V上每个线性函数都形如$\mathbf{y} \mapsto (\mathbf{x}|\mathbf{y})$, $\mathbf{x}$是某个固定向量. 在这种联系中, 我们可回想对于对偶基底的表达式$e^i = (\mathbf{e}_i|*)$(见§1的(13)式).

值得指出, 在没有附加欧几里得结构或者埃尔米特结构的向量空间中, 在某个基底之下, 把θ线性型f与F, 线性算子与矩阵$A = {}^tF$相对应是带有偶然性的. 事实上, 借助转换矩阵B向新基底转换时, θ线性型的矩阵就变成$F' = {}^tBF\overline{B}$, 因此, $A' = {}^tF' = B^{*t}FB$. 但是, 与此同时, 按第2章§2的定理3, 我们应当有$A' = B^{-1}AB = B^{-1t}FB$. 这两个对于$A'$的表达式没有共同之处.但在埃尔米特(欧几里得)空间情形, 矩阵B应该保持基底的标准正交性, 这就导致它的酉性(相应地, 正交性). 对于酉矩阵$B^* = B^{-1}$, 就有了完全一致性.

再一次把映射$\mathcal{A} \mapsto \mathcal{A}^*$的性质写下来:

$$(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, \quad (\alpha\mathcal{A})^* = \overline{\alpha}\mathcal{A}^*, \quad (\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*, \quad \mathcal{A}^{**} = \mathcal{A}. \tag{5}$$

与第2章§3的公式(15)一个不太大的差别在于, 在α上有一个复共轭标识, 这是由型$(*|*)$的半双线性性质和对$\mathbf{\Phi}_\mathbf{v}: f \mapsto (\mathbf{v}|*)$的类型引起的(§1,第4目).

2. 线性算子的类型 作用在一个带有纯量乘积$(*|*)$的向量空间V上的所有的线性算子可按§1引入的对于算子$*$的关系分成与它们状态有关系的类. 我们只标出一些极重要的类.

定义1 称线性算子$\mathcal{A}$是**埃尔米特的** (或**自共轭的**), 如果$\mathcal{A}^* = \mathcal{A}$. 在欧几里得空间情形($\mathfrak{K} = \mathbb{R}$), 算子$\mathcal{A} = \mathcal{A}^*$还称为是**对称的**.

正如定理表明的那样, 算子$\mathcal{A}$的自对称性等价于θ线性型$(\mathcal{A}\mathbf{x}|\mathbf{y})$的埃尔米特性质. 事实上, 自共轭性的条件可以记成

$$(\mathcal{A}\mathbf{x}|\mathbf{y}) = (\mathbf{x}|\mathcal{A}\mathbf{y}),$$

而型$f_{\mathcal{A}}$的埃尔米特性质可记为

$$(\mathcal{A}\mathbf{x}|\mathbf{y}) = f_{\mathcal{A}}(\mathbf{x},\mathbf{y}) = \overline{f_{\mathcal{A}}(\mathbf{y},\mathbf{x})} = \overline{(\mathcal{A}\mathbf{y}|\mathbf{x})}.$$

因为$(*|*)$是埃尔米特型, 所以$\overline{(\mathcal{A}\mathbf{y}|\mathbf{x})} = (\mathbf{x}|\mathcal{A}\mathbf{y})$. 这就建立了所提到的两个条件的等价性.

在矩阵形式中, 如果利用空间V的标准正交基底, 算子$\mathcal{A}$的自共轭性(埃尔米特性)条件可用等式${}^t\overline{A} - A = 0$表示出来.前面, 我们已经把这样的矩阵称为埃尔米特矩阵了. 而在实的情形, 就是对称矩阵.

每个实矩阵都是一个对称矩阵和一个斜对称矩阵的和(见第1章§4的第4目).为了在复数情形能得到类似的结果, 我们引进

定义2　斜线性算子是**斜埃尔米特的**(或者, 当$\mathfrak{K} = \mathbb{R}$时, **斜对称的**), 如果

$$\mathcal{A}^* = -\mathcal{A}.$$

因为, 对任意$\mathcal{A} \in \mathcal{L}(V)$, 有$\mathcal{A}^{**} = \mathcal{A}$, 所以$\mathcal{A} + \mathcal{A}^*$是埃尔米特的, 而$\mathcal{A} - \mathcal{A}^*$则是斜埃尔米特的. 类似地, $\mathcal{A}$的埃尔米特性质等价于算子$i\mathcal{A}$的斜埃尔米特性质. 因此, 有

定理2　在埃尔米特空间中, 每个线性算子$\mathcal{Z}$都可以写成

$$\mathcal{Z} = \mathcal{A} + \mathcal{B}$$

的形式, 其中$\mathcal{A}$是埃尔米特的, 而$\mathcal{B}$是个斜埃尔米特算子. 此外

$$\mathcal{Z} = \mathcal{X} + i\mathcal{Y}, \tag{6}$$

其中$\mathcal{X}$和$\mathcal{Y}$是埃尔米特线性算子.

证明　令$\mathcal{A} = (\mathcal{Z} + \mathcal{Z}^*)/2$, $\mathcal{B} = (\mathcal{Z} - \mathcal{Z}^*)/2$, $\mathcal{X} = \mathcal{A}$, $\mathcal{Y} = -i\mathcal{B}$.然后, 借助公式(5)即可直接验证. □

显然, 记法(6)是把复数$z = x + iy$的记法的直接推广, 就是说, 埃尔米特算子有些类似于实数. 而说到斜埃尔米特算子本身就是虚数$z = iy$的直接“后代”数, 虚数有性质, $\overline{z} = -z$. 但是, 如果说, 任意两个实数之积永远是实数, 那么, 两个埃尔米特算子的乘积却不一定仍然是埃尔米特算子.

我们总有

定理3　埃尔米特算子$\mathcal{A}, \mathcal{B}$的乘积$\mathcal{AB}$是埃尔米特的, 当且仅当, $\mathcal{AB} = \mathcal{BA}$.

证明　再次利用(5), 就得到

$$\mathcal{AB} = \mathcal{BA} \Leftrightarrow (\mathcal{AB})^* = (\mathcal{BA})^* = \mathcal{A}^*\mathcal{B}^* = \mathcal{AB}.$$ □

数学和物理学中大量的分解问题都要求把所有的埃尔米特算子的集合或者所有的斜埃尔米特算子的集合在第2章§2定义1的意义下作为一个代数来进行研究.

例1 正如在定理3中看到的, 一般说来, 埃尔米特矩阵或算子对结合乘法并不一定是封闭的. 为了寻求量子力学的代数学框架, 物理学家P. 若尔当于1930年在$\mathbb{R}$上引入一种代数, 现称为若尔当代数. 作为基数, 规定了若尔当乘积

$$\mathcal{A}\circ\mathcal{B}=\frac{1}{2}\left(\mathcal{A}\mathcal{B}+\mathcal{B}\mathcal{A}\right),$$

显然它满足交换律, 也满足若尔当恒等式$(\mathcal{A}^2\circ\mathcal{B})\circ\mathcal{A}=\mathcal{A}^2\circ(\mathcal{B}\circ\mathcal{A})$(验证之!). 到现在, 内容丰富多彩的若尔当代数理论(不仅仅限于有限维)已经很发达了.

例2 所有的斜埃尔米特算子对于通常的换位子运算构成了$\mathbb{R}$上的一个李代数(参见第2章§2的例6). 换言之, 如果$\mathcal{A}$和$\mathcal{B}$都是斜埃尔米特算子, 那么, 它们的换位子$[\mathcal{A},\mathcal{B}]=\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A}$同样是个斜埃尔米特算子.

我们设, 对任意$\mathbf{x}$, $\mathbf{y}\in V$都有$(\mathcal{A}\mathbf{x}|\mathbf{y})=0$. 那么, 特别地, $(\mathcal{A}\mathbf{x}|\mathcal{A}\mathbf{x})=0$. 但是, 这只有对任意$\mathbf{x}\in V$, $\mathcal{A}\mathbf{x}=\mathbf{0}$才行, 也就是说, $\mathcal{A}=\mathcal{O}$. 这个关于$\mathcal{A}$的平凡性的判别法实际上还可以加强.

定理4 *设, 对任意$\mathbf{x}\in V$都有$(\mathcal{A}\mathbf{x}|\mathbf{x})=0$, 而且$\mathcal{A}$满足下列二条件之一:*

i) V是个埃尔米特空间;

ii) V是个欧几里得空间, 且$\mathcal{A}$是对称算子.

那么, $\mathcal{A}=\mathcal{O}$.

证明 i) 由下面两个很容易验证的极化恒等式

$$(\mathcal{A}\mathbf{x}|\mathbf{y})+(\mathcal{A}\mathbf{y}|\mathbf{x})=(\mathcal{A}\left(\mathbf{x}+\mathbf{y}\right)|\mathbf{x}+\mathbf{y})-(\mathcal{A}\mathbf{x}|\mathbf{x})-(\mathcal{A}\mathbf{y}|\mathbf{y}),\tag{7}$$

$$(\mathcal{A}\mathbf{x}|\mathbf{y})-(\mathcal{A}\mathbf{y}|\mathbf{x})=-i\left(\mathcal{A}\left(i\mathbf{x}+\mathbf{y}\right)|i\mathbf{x}+\mathbf{y}\right)+i\left(\mathcal{A}\left(i\mathbf{x}\right)|i\mathbf{x}\right)+i\left(\mathcal{A}\mathbf{y}|\mathbf{y}\right)\tag{8}$$

(按假设, 它们的右侧应该等于0), 我们就可以推导出两个齐次线性方程构成的方程组

$$(\mathcal{A}\mathbf{x}|\mathbf{y})+(\mathcal{A}\mathbf{y}|\mathbf{x})=0,\quad(\mathcal{A}\mathbf{x}|\mathbf{y})-(\mathcal{A}\mathbf{y}|\mathbf{x})=0.$$

由此可得$(\mathcal{A}\mathbf{x}|\mathbf{y})=0$, $\forall\mathbf{x},\mathbf{y}\in V$, 而这正是我们已经知道的, 它等价于$\mathcal{A}=\mathcal{O}$.

ii) 任意情形都满足的极化恒等式(7)和对称条件

$$(\mathcal{A}\mathbf{y}|\mathbf{x})=(\mathbf{y}|\mathcal{A}^*\mathbf{x})=(\mathbf{y}|\mathcal{A}\mathbf{x})=(\mathcal{A}\mathbf{x}|\mathbf{y})$$

给出了, 任意$\mathbf{x},\mathbf{y}\in V$都满足恒等式$(\mathcal{A}\mathbf{x}|\mathbf{y})=0$本身, 由此得出, $\mathcal{A}=\mathcal{O}$. □

说明1 在定理4条件ii)中, 算子$\mathcal{A}$的对称性是本质的. 例如, 对于欧几里得空间中的斜对称算子, 恒等式$(\mathcal{A}\mathbf{x}|\mathbf{x})=0$是对任意$\mathbf{x}\in V$都满足的, 但$\mathcal{A}$并不一定就是零算子.

定义3 *在带有纯量乘积的向量空间V上, 称线性算子$\mathcal{A}$为***酉算子***(在欧几里得空间情形是***正交算子***), 如果$\mathcal{A}^*\cdot\mathcal{A}=\mathcal{E}=\mathcal{A}\cdot\mathcal{A}^*$.*

当n=1时, 有$z\cdot\overline{z}=1$, 此时酉算子就类似于复数的模等于单位元. 在矩阵形式(相对于标准正交基底)酉性条件可由§2的等式(7)表达出来. 从而, 那样的矩阵, 我们也

称它为酉矩阵(在实的情形即正交矩阵). 它们是作为从一个标准正交基底向另一个标准正交基底的转换矩阵用极其自然的方法产生的, 这个事实回报了酉算子的更丰富多彩的几何解释.

定义4 线性算子$\mathcal{A}: V \to V$保持距离(度量), 也就是, 有

$$\|\mathcal{A}\mathbf{x} - \mathcal{A}\mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in V,$$

即被称为是**保距**的.

因为, $\mathcal{A}\mathbf{x} - \mathcal{A}\mathbf{y} = \mathcal{A}(\mathbf{x} - \mathbf{y})$, 所以, 显然, $\mathcal{A}$是保距的, 当且仅当$\|\mathcal{A}\mathbf{x}\| = \|\mathbf{x}\|$对所有$\mathbf{x} \in V$都成立. 其次

$$\begin{aligned} \|\mathcal{A}\mathbf{x}\| = \|\mathbf{x}\| &\Leftrightarrow (\mathcal{A}\mathbf{x}|\mathcal{A}\mathbf{x}) = (\mathbf{x}|\mathbf{x}) \\ &\Leftrightarrow (\mathcal{A}^*\mathcal{A}\mathbf{x}|\mathbf{x}) = (\mathbf{x}|\mathbf{x}) \Leftrightarrow ((\mathcal{A}^*\mathcal{A} - \mathcal{E})\,\mathbf{x}|\mathbf{x}) = 0 \end{aligned} \tag{9}$$

对任何$\mathbf{x} \in V$都成立.算子$\mathcal{A}^*\mathcal{A} - \mathcal{E}$是自共轭的, 所以, 据定理4, 作为一个埃尔米特算子, 同样也在欧几里得情形, 由(9)可导出恒等式$\mathcal{A}^*\mathcal{A} - \mathcal{E} = \mathcal{O}$, 即, 保距算子应当是酉算子.

另一方面, 所有的酉算子都是保距的:

$$(\mathcal{A}\mathbf{x}|\mathcal{A}\mathbf{x}) = (\mathbf{x}|\mathcal{A}^*\mathcal{A}\mathbf{x}) = (\mathbf{x}|\mathcal{E}\mathbf{x}) = (\mathbf{x}|\mathbf{x}).$$

从而验证了

定理5 在附有度量的向量空间V上, 酉线性算子是V上的保度量算子, 而且只有它们才是保度量的.

由此可见, V上的酉算子, 也就是保度量算子, 构成一个群——当$\mathfrak{K} = \mathbb{C}$时, 是**酉群**$U(n)$, 而当$\mathfrak{K} = \mathbb{R}$时, 就是**正交群**$O(n)$. 用矩阵语言, 这些事实我们已经熟悉了(§2的定理2). 在这里, 群$GL(n, \mathfrak{K})$处于同样地位: 可以讨论矩阵群, 也可以讨论空间V的所有自同构的群$\mathrm{Aut}(V)$. 保距矩阵就是保持度量的自同构.

3. 埃尔米特算子的规范形式 对任意一个埃尔米特算子都存在一个由它的特征向量组成的基底, 乍眼一看这并不是一个明显的性质. 事实上, 给定的对称算子$\mathcal{A}: V \to V$(当$\mathfrak{K} = \mathbb{R}$时)在不同的标准正交基底$(\mathbf{e}_i)$, $(\mathbf{e}'_i)$之下对应的矩阵A, A'可用关系式$A' = B^{-1}AB$联系起来, 其中B是个正交矩阵. 我们知道, 对称矩阵可以化成对角形式, 关键是这个非退化矩阵选择的随意性. 正如已经表明的那样, $\mathcal{A}$的自对偶性质可以更"经济"地加以支配.

引理1 埃尔米特算子的特征值(固有值)都是实的.

证明 事实上, 设$\mathcal{A}: V \to V$是个埃尔米特算子, λ是它的一个与特征向量$\mathbf{e}$相适应的特征值. 按定义,

$$\lambda(\mathbf{e}|\mathbf{e}) = (\lambda\mathbf{e}|\mathbf{e}) = (\mathcal{A}\mathbf{e}|\mathbf{e}) = (\mathbf{e}|\mathcal{A}^*\mathbf{e}) = (\mathbf{e}|\mathcal{A}\mathbf{e}) = (\mathbf{e}|\lambda\mathbf{e}) = \overline{\lambda}(\mathbf{e}|\mathbf{e}).$$

因为$(\mathbf{e}|\mathbf{e}) \neq 0$, 所以$\overline{\lambda} = \lambda$. □

在对称的(即实的自共轭的)算子情形, 引理1是个空命题, 因为所有的特征值按定义都应属于$\mathbb{R}$, 反过来, 在复的情形, 下面的引理是显然的.

引理2 每个对称的线性算子$\mathcal{A}$都必有特征向量.

证明 作为任意一个实的算子, $\mathcal{A}$必有1维或者2维的特征子空间(第2章§3的定理7).1维不变子空间的存在性与本引理的断言是一回事. 现在研究L是2维不变子空间的情形. 算子$\mathcal{A}$在L上诱导出一个对称的线性算子$\mathcal{A}_L$. 这是因为, 对称性条件$(\mathcal{A}\mathbf{x}|\mathbf{y}) = (\mathbf{x}|\mathcal{A}\mathbf{y})$当作$\mathbf{x}, \mathbf{y} \in L$上的限制时可继续保持正确: $\mathcal{A}\mathbf{x} \in L$, $\mathcal{A}\mathbf{y} \in L$.

在L中选择标准正交基底$(\mathbf{e}_1, \mathbf{e}_2)$. 算子$\mathcal{A}_L$在这个基底之下的矩阵就是一个$2 \times 2$的对称矩阵

$$A_L = \begin{pmatrix} a & b \\ b & d \end{pmatrix},$$

它的特征多项式是

$$\chi(t) = \begin{vmatrix} t-a & -b \\ -b & t-d \end{vmatrix} = t^2 - (a+d)\,t + \left(ad - b^2\right).$$

这个多项式的判别式

$$D_\chi = (a+d)^2 - 4\left(ad - b^2\right) = (a-d)^2 + 4b^2 \geqslant 0,$$

所以$\chi(t)$必有实根λ, 从而算子$\mathcal{A}$就有属于特征值λ的特征向量. □

可同时对$\mathfrak{K} = \mathbb{C}$和$\mathfrak{K} = \mathbb{R}$推导出更进一步的论断.

引理3 设$\mathcal{A}$是带有纯量乘积$(*|*)$的向量空间V上的一个自共轭线性算子, L是个对$\mathcal{A}$不变的子空间. 那么, L的正交补空间$L^\perp$也是对$\mathcal{A}$不变的子空间.

证明 事实上, 如果$\mathbf{x} \in L$, $\mathbf{y} \in L^\perp$,那么$\mathcal{A}\mathbf{x} \in L$且$(\mathcal{A}\mathbf{x}|\mathbf{y}) = 0$. 算子$\mathcal{A}$的自共轭性给出同样的关系式$(\mathbf{x}|\mathcal{A}\mathbf{y}) = 0$. 由此可见, $\mathcal{A}\mathbf{y}$与所有的向量$\mathbf{x} \in L$都正交, 即$\mathcal{A}L^\perp \subset L^\perp$. □

现在, 我们准备动手证明基本定理.

定理6 带有纯量乘积的向量空间V必有标准正交基底, 在这个基底之下, 自共轭算子$\mathcal{A}$的矩阵是对角的, 而且$\mathrm{Spec}(\mathcal{A})$是个实数集.

证明 根据引理1和引理2, 线性算子$\mathcal{A}$必有属于特征值$\lambda_1 \in \mathbb{R}$的特征向量$\mathbf{e}_1$. 不失一般性, 可以认为$\|\mathbf{e}_1\| = 1$. 1维子空间$\langle \mathbf{e}_1 \rangle$的正交补空间$V'$的维数是$\dim V - 1$, 且据引理3, 它对$\mathcal{A}$也是不变的. 看$\mathcal{A}$在$V'$上的限制并重复所有的推断, 可找出$\mathbf{e}_2$: $\mathcal{A}\mathbf{e}_2 = \lambda_2 \mathbf{e}_2$, $\|\mathbf{e}_2\| = 1$, $\lambda_2 \in \mathbb{R}$. 线性包络$\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$对于$\mathcal{A}$也是不变的, 因此. 和它正交的补空间也是对$\mathcal{A}$不变的, 维数为$\dim V - 2$, 等等. 对$\dim V$用归纳法, 或者简单地重复应有数量的前述过程, 我们就得到了所需要的$n = \dim V$个相互正交的标准的向量$\mathbf{e}_1, \cdots, \mathbf{e}_n$. □

说明2 任意对称矩阵$A \in M_n(\mathbb{R})$的特征方程式已被证明只有实根. 要研究它们间的相互位置, 可用 [BA I] 中的笛卡儿, 比丹–傅里叶, 斯图姆定理. 方程式$\chi_A(t) = 0$的每个根λ的几何重数和代数重数相同, 这可以由第2章§3的定理6和本节定理6直接得到.

说明3 按定理6, 对任意自共轭算子$\mathcal{A}: V \to V$有$n = \dim V$个相互正交的特征方向. 算子$\mathcal{A}$的作用可以归结为空间沿第k个方向延长$|\lambda_k|$倍, 其中λ_k是相应的特征值, 而且可能, 当$\lambda_k < 0$时, 归结为一个与第k个方向正交的平面反射.

4. 把二次型化到主轴上去 我们知道(见第1目), 在带有纯量乘积$(*|*)$的向量空间 V上, 每个埃尔米特型$f(\mathbf{x}, \mathbf{y})$都对应一个算子$\mathcal{A} = \mathcal{A}_f$, 它由条件

$$f(\mathbf{x}, \mathbf{y}) = (\mathcal{A}\mathbf{x}|\mathbf{y})$$

决定. 据定理6, 空间 V有由算子$\mathcal{A}$的特征向量组成的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$,$\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$. 如果把向量$\mathbf{x}, \mathbf{y}$在这个坐标系下写出来:

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n, \quad \mathbf{y} = y_1\mathbf{e}_1 + \cdots + y_n\mathbf{e}_n,$$

那么, 我们就得到

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j} f(\mathbf{e}_i, \mathbf{e}_j) x_i \overline{y_j} = \sum_{i=1}^{n} \lambda_i x_i \overline{y_i},$$

这是因为$f(\mathbf{e}_i, \mathbf{e}_j) = (\mathcal{A}\mathbf{e}_i|\mathbf{e}_j) = (\lambda_i \mathbf{e}_i|\mathbf{e}_j) = \lambda_i \delta_{ij}$.设$\mathbf{x} = \mathbf{y}$, 我们就导出下列命题

定理7(化到主轴上) 对于带有纯量乘积的n维向量空间上的每个埃尔米特二次型$q(\mathbf{x})$都存在一个标准正交基底, 在这个基底之下, $q(\mathbf{x})$形如

$$q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i |x_i|^2. \tag{9'}$$

例3 当$\mathfrak{K} = \mathbb{R}$且$n$=2时, 二次型确定中心二次曲线截口, 它由使得$q(\mathbf{x}) = 1$的向量$\mathbf{x}$组成. 利用标准正交基底$(\mathbf{e}_1, \mathbf{e}_2)$, 在这个基底之下$q(\mathbf{x})$形如(9′)式, 有$\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$. 向量$\mathbf{e}_1, \mathbf{e}_2$定义了椭圆$(\lambda_1\lambda_2 > 0)$或者双曲线$(\lambda_1\lambda_2 < 0)$的主轴方向, 而用$\lambda_1, \lambda_2$代表半轴长度.

定理6和定理7同样可以用矩阵语言表达出来.

对任意埃尔米特(或实对称)矩阵A都有酉矩阵(相应地, 正交矩阵)B使得$B^{-1}AB$是对角的. 沿对角线排列A的特征值, 按特征值的重数排列它的个数.

可实现性简介 几何事实的矩阵解释提示了将二次型

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ij} x_i x_j$$

(限于实的情形)化为标准的作法的可能顺序.也就是, 在带有纯量乘积

$$(\mathbf{x}|\mathbf{y}) = \sum_{i=1}^{n} x_i y_i,$$

的欧几里得空间 V中, 以$x_1, \cdots, x_n$为向量$\mathbf{x}$的坐标, 而$\mathbf{e}_1 = (1,0,\cdots,0), \cdots, \mathbf{e}_n = (0,0,\cdots,1)$为 V的标准正交基底. 计算矩阵$A = (a_{ij})$的特征多项式$\chi_A(t) = \det(tE-A)$且找出它的根(这是整个过程中的最困难部分).对每个根λ_i解齐次线性方程组

$$\begin{aligned}
&(a_{11} - \lambda_i) x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0,\\
&a_{21}x_1 + (a_{22} - \lambda_i) x_2 + \cdots + a_{2n}x_n = 0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda_i) x_n = 0.
\end{aligned}$$

这个方程组的解空间的维数等于根λ_i的代数重数(矩阵A的对称性的直接推论). 把格拉姆–施密特正交化过程用到方程组的基础解系上, 然后再把不同的λ_i所对应的向量合并起来, 我们就得到 V的一个标准正交基底

$$\mathbf{e}'_j = b_{1j}\mathbf{e}_1 + b_{2j}\mathbf{e}_2 + \cdots + b_{nj}\mathbf{e}_n, \quad 1 \leqslant j \leqslant n.$$

在这里, 就其本质来说, 我们依靠了算子$\mathcal{A}$及其在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$下的矩阵$A = (a_{ij})$的对称性(更一般地, 自共轭性): 分别属于不同特征值λ, μ的特征向量$\mathbf{u}$, $\mathbf{v}$相互正交. 事实上, $(\mathcal{A}\mathbf{u}|\mathbf{v}) = (\mathbf{u}|\mathcal{A}\mathbf{v}) \Rightarrow (\lambda\mathbf{u}|\mathbf{v}) = (\mathbf{u}|\mu\mathbf{v}) \Rightarrow (\lambda - \mu)(\mathbf{u}|\mathbf{v}) = 0 \Rightarrow (\mathbf{u}|\mathbf{v}) = 0$(重要的是$\lambda$, μ均为实数).

联系两个标准正交基底的矩阵(b_{ij})一定是正交的(在我们这里, $\mathfrak{K} = \mathbb{R}$).所以, 向量$\mathbf{x}$的新的坐标$x'_1, \cdots, x'_n$可以按§1末尾的公式用原来的坐标表达出来, 而且

$$q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i (x'_i)^2 .$$

5. 把两个二次型同时化为规范型 以二次型$q(\mathbf{x}) = |x_1|^2 - |x_2|^2$, $r(\mathbf{x}) = |\mathbf{x}_1| \cdot |\mathbf{x}_2|$为例, 可以表明, 并不总是能够在向量空间中选取一个基底, 使得两个二次型同时化成规范型.但是, 在一种可实现的重要情况, 上面提到的基底的存在性还是有保障的.

定理8 设在$\mathfrak{K} = \mathbb{R}$或者$\mathfrak{K} = \mathbb{C}$上的$n$维向量空间$V$上给定了两个埃尔米特二次型(即实二次型)$q(\mathbf{x})$和$r(\mathbf{x})$, 其中$r(\mathbf{x})$是正定的.

那么, 在V中存在一个基底, 在该基底之下, 两个二次型都可写成规范形式.

证明 设$g(\mathbf{x}, \mathbf{y})$是二次型$r(\mathbf{x})$对应的埃尔米特$\theta$线性型. 我们在 V上定义纯量乘积, 令

$$(\mathbf{x}|\mathbf{y}) := g(\mathbf{x}, \mathbf{y}) .$$

二次型$r(\mathbf{x})$的正定性允许这样做. 按定理7, 在V中对一个给定的埃尔米特矩阵可找到一个标准正交基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$, 在这个基底之下$q(\mathbf{x})$具有形如(9)的规范型.

于是, 在这个同样的基底之下按公式

$$(\mathbf{x}|\mathbf{x}) = g(\mathbf{x},\mathbf{x}) = r(\mathbf{x}) = \sum_{i=1}^{n} |x_i|^2$$

可计算出来纯量二次型.也就是说, 在基底$(\mathbf{e}_i)$之下, 两个二次型都具规范型. □

6. 保距算子的规范形式 依据定理5, V上的保距算子, 精确地说就是一个酉算子(当$\mathfrak{K}=\mathbb{R}$时就是正交算子).我们分复的情形和实的情形分别加以研究.但是先从证明一些一般性事实开始.

引理4 酉(正交)算子的特征值的模等于1(实的情形, 特征值为±1).

证明 设$\mathcal{A}:V\to V$是个酉算子(特别地, 正交算子), 且$\mathbf{e}\in V$是属于特征值λ的特征向量.那么,

$$(\mathcal{A}\mathbf{e}|\mathcal{A}\mathbf{e}) = (\lambda\mathbf{e}|\lambda\mathbf{e}) = \lambda\overline{\lambda}.$$

另一方面,

$$(\mathcal{A}\mathbf{e}|\mathcal{A}\mathbf{e}) = (\mathcal{A}^*\cdot\mathcal{A}\mathbf{e}|\mathbf{e}) = (\mathcal{E}\mathbf{e}|\mathbf{e}) = (\mathbf{e}|\mathbf{e}).$$

所以, $\lambda\overline{\lambda}=1$, 也就是$|\lambda|=1$. 显然, 在实的情形(正交算子)只有$\lambda=\pm1$两个可能. □

引理5 设$U\subset V$是个酉(正交)算子$\mathcal{A}:V\to V$的一个不变子空间. 那么, U在V中的正交补$U^{\perp}$对于$\mathcal{A}$也同样是不变的.

证明 按定义

$$U^{\perp} = \{\mathbf{v}\in V\,|\,(\mathbf{u}|\mathbf{v})=0 \quad \forall\mathbf{u}\in U\}.$$

算子$\mathcal{A}$在U上的限制$\mathcal{A}_U$显然是个酉算子(在U上是保距的). 因为$\det\,\mathcal{A}_U\neq0$, 所以向量$\mathbf{u}\in U$可以写成$\mathbf{u}=\mathcal{A}\mathbf{u}'$, $\mathbf{u}'\in U$. 我们有

$$(\mathbf{u}|\mathcal{A}\mathbf{v}) = (\mathcal{A}\mathbf{u}'|\mathcal{A}\mathbf{v}) = (\mathbf{u}'|\mathbf{v}) = 0.$$

换句话说, 只要$\mathbf{v}\in U^{\perp}$, 就必有$\mathcal{A}\mathbf{v}\in U^{\perp}$. □

I. 酉算子情形 用矩阵术语, 我们想证明下面的命题: 对每个酉矩阵A都存在一个酉矩阵B使得

$$C = B^{-1}AB = \mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}$$

是个对角矩阵, 且$|\lambda_i|=1$.

实际上, 借助算子的几何意义, 运作起来更方便.

定理9 每个酉算子($\mathfrak{K}=\mathbb{C}$)都可以对角化.换言之, 对任意一个酉算子$\mathcal{A}:V\to V$, $\dim\,V=n$, 都可以找到一个标准正交基底, 在这个基底之下, 算子的矩阵是

$$A = \mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}, \quad |\lambda_i| = 1. \tag{10}$$

证明 任取算子$\mathcal{A}$的一个标准的特征向量e_1. 它一定存在, 因为基础域$\mathfrak{K} = \mathbb{C}$是个代数封闭域. 由引理5, 维数为$n-1$的子空间$\langle \mathrm{e}_1 \rangle^{\perp} = U$对$\mathcal{A}$是不变的. 对$V$的维数用归纳法就可得到所需的结论. 关于$\lambda_i$的断言已经在引理4中证明过了. □

注意, 具有形如(10)的矩阵A的算子, 显然是个酉算子, 实际上

$$^t\overline{A} \cdot A = \mathrm{diag}\left\{\overline{\lambda_1}, \cdots, \overline{\lambda_n}\right\} \cdot \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_n\right\} = E.$$

通常, 利用欧拉公式$e^{i\varphi} = \cos\ \varphi + i\ \sin\ \varphi$, 将对角形矩阵记成

$$A = \begin{pmatrix} e^{i\varphi_1} & & & \\ & e^{i\varphi_2} & & \\ & & \ddots & \\ & & & e^{i\varphi_n} \end{pmatrix}$$

形式.

Ⅱ. 正交算子情形 正如已经指出过的, 引理4和引理5对于正交算子$\mathcal{A}: V \to V$也是正确的, 而且已经特别声明, 在这种情形, 特征值等于±1. 但是, 将要进行的推理仍然需要做不少的变换. 事情在于, 正交变换可能根本就没有特征向量. 当然, 作为任意一个实的线性算子, $\mathcal{A}$必定有1维的或者2维的不变子空间. 因此, 我们可以借助引理5把V分解成两两正交的1维或者2维的不变子空间的直和:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m, \tag{11}$$

$\mathcal{A}$在它们之中的每一个上面都导出一个正交的线性算子. 把$V_i, i = 1, 2, \cdots, m$的标准正交基底合并到一起, 我们就得到了V的一个标准正交基底. 假设在(11)的分解中, 没有任何一个2维的不变子空间能够分解成两个1维子空间的直和, 这就得到了正交线性算子$\mathcal{A}$的一个所谓的**规范基底**. 这一点, 通过努力是一定能达到的, 所以, 我们假定分解式(11)就已经是这样的了.

现在研究在规范基底之下算子$\mathcal{A}$的矩阵是怎样的. 如果, $A_i := A_{V_i}$是算子$\mathcal{A}$在V_i上限制算子的矩阵, 那么,

$$A = A_1 \dot{+} \cdots \dot{+} A_m = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}.$$

所以, 我们限制在$\dim V = 1$或者$\dim V = 2$而且V没有非平凡的不变子空间这样的条件下就足够了. 如果$V = \langle \mathrm{e} \rangle$, $\|\mathrm{e}\| = 1$, 那么$\mathcal{A}\mathrm{e} = \lambda \mathrm{e}$, $\lambda = \pm 1$ (引理4). 如

果$\dim V=2$, $V=\langle \mathbf{e}_1,\mathbf{e}_2\rangle$, 且$(\mathbf{e}_i|\mathbf{e}_j)=\delta_{ij}$, 那么, 在这个标准正交基底之下

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix},\quad a,b,c,d\in\mathbb{R}.$$

我们假定$\det A=ad-bc=-1$, 那么, 特征多项式$\chi_A(t)=t^2-(a+d)t-1$有两个实根, 可见, 算子$\mathcal{A}$有特征向量, 但这与前面在V上的假设的条件相矛盾. 于是得出结论, $\det A=1$. 按已知的规则把逆矩阵A^{-1}算出来, 得

$$A^{-1}=\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

另一方面, 由正交标准性

$$A^{-1}={}^tA=\begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

比较A^{-1}的这两个表达式, 我们有

$$A=\begin{pmatrix} a & -c \\ c & a \end{pmatrix},\quad a^2+c^2=1.$$

这样一来, 当$a=\cos\varphi$, $c=\sin\varphi$时, 就有

$$A=\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

也就是说, 线性算子$\mathcal{A}$实现了平面V上的一个旋转.

上面所做的分析表明, 假设分解式(11)中前面的r个被加项$V_1,\cdots,V_r$对应2维的不可分解的不变子空间, 而剩下的被加项对应1维不变子空间(经过调换基底中向量的顺序, 这是一定能达到的), 再设$\varphi_1,\cdots,\varphi_r$是这些旋转分别对应的角度, 那么, 矩阵$A$具有下面定理指明的形式.

定理10 对V上的每个的正交线性算子$\mathcal{A}$都存在一个V的标准正交基底, 在这个基底之下, 算子的矩阵是

$$\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 & & & & & \\ \sin\varphi_1 & \cos\varphi_1 & & & & & \\ & & \ddots & & & & \\ & & & \cos\varphi_r & -\sin\varphi_r & & \\ & & & \sin\varphi_r & \cos\varphi_r & & \\ & & & & & -E_k & \\ & & & & & & E_l \end{pmatrix},\quad k+l+2r=n.$$

7. 正规算子　在已证实的谱定理, 定理6和定理9中有很多共同之处, 而这不是偶然的, 因为可以把埃尔米特算子和酉算子加入到自然的, 更广泛的可对角化的矩阵的行列中去.

定义5　设V是个埃尔米特空间.称线性算子$\mathcal{A}:V\to V$是**正规的**, 如果它具有性质

$$\mathcal{A}\cdot\mathcal{A}^*=\mathcal{A}^*\cdot\mathcal{A}. \tag{12}$$

它在任意一个基底之下的矩阵都称为**正规矩阵**.

回想一下, 由(5)有关系式

$$(\lambda\mathcal{E})^*=\overline{\lambda}\mathcal{E},\quad (\mathcal{A}-\lambda\mathcal{E})^*=\mathcal{A}^*-\overline{\lambda}\mathcal{E},$$

所以, 算子$\mathcal{A}$是正规的, 当且仅当, $\mathcal{A}-\lambda\mathcal{E}$是正规的. 由$\mathcal{A}$的正规性得出

$$\|\mathcal{A}\mathbf{x}\|^2=(\mathcal{A}\mathbf{x}|\mathcal{A}\mathbf{x})=(\mathbf{x}|\mathcal{A}^*\mathcal{A}\mathbf{x})=(\mathbf{x}|\mathcal{A}\mathcal{A}^*\mathbf{x})=(\mathcal{A}^*\mathbf{x}|\mathcal{A}^*\mathbf{x})=\|\mathcal{A}^*\mathbf{x}\|^2.$$

用$\mathcal{A}-\lambda\mathcal{E}$替换$\mathcal{A}$, 我们得到

$$\|\mathcal{A}\mathbf{x}-\lambda\mathbf{x}\|=\left\|\mathcal{A}^*\mathbf{x}-\overline{\lambda}\mathbf{x}\right\|,$$

由此推出

$$\mathcal{A}\mathbf{x}=\lambda\mathbf{x}\Leftrightarrow\mathcal{A}^*\mathbf{x}=\overline{\lambda}\mathbf{x}. \tag{13}$$

显然, 由条件$\mathcal{A}^*=\mathcal{A}$或者$\mathcal{A}^*=\mathcal{A}^{-1}$都能导致(12)成立. 但是, 一点也不难就可以举出正规算子的例子, 它既不是埃尔米特(或斜埃尔米特)算子, 也不是酉算子(比方说, 取矩阵$A=\operatorname{diag}(2i,2,1,\cdots 1)$). 同时, 正规算子的定义可以搬到无穷维希尔伯特空间并且在那里有大量的应用. 我们的直接目标是在埃尔米特空间上精确地刻画可以对角化的算子类.

定理11　下述条件是等价的:

i) 算子$\mathcal{A}:V\to V$在V的某个标准正交基底之下是对角的;

ii) 算子$\mathcal{A}$是正规的.

证明　i)⇒ii). 如果$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是标准正交的, 且$\mathcal{A}\mathbf{e}_i=\lambda_i\mathbf{e}_i$, 那么, 借助(13), 有 $\mathcal{A}^*\mathbf{e}_i=\overline{\lambda}_i\mathbf{e}_i$, 故$[\mathcal{A},\mathcal{A}^*]=\mathcal{O}$, 就是说, 可由i)得到ii).

为了证明反方向的蕴涵式ii)⇒ i), 我们选出$\mathcal{A}$的特征值λ且照例设

$$V^\lambda=\{\mathbf{x}\in V|\mathcal{A}\mathbf{x}=\lambda\mathbf{x}\}.$$

再由(13)得到

$$\mathcal{A}^*\left(V^\lambda\right)\subseteq V^\lambda,$$

而且, 同样有,

$$\mathcal{A}\left(V^\lambda\right)^\perp\subseteq\left(V^\lambda\right)^\perp.$$

事实上,

$$\mathbf{y} \in \left(V^{\lambda}\right)^{\perp} \Leftrightarrow (\mathbf{y}|\mathbf{x}) = 0, \quad \forall \mathbf{x} \in V^{\lambda}.$$

由此可见, $(\mathcal{A}\mathbf{y}|\mathbf{x}) = (\mathbf{y}|\mathcal{A}^*\mathbf{x}) = (\mathbf{y}|\mathbf{x}') = 0$, 因为$\mathbf{x}' \in V^{\lambda}$.

因为$(\mathcal{A}^*)^* = \mathcal{A}$, 所以由对称性, 子空间$\left(V^{\lambda}\right)^{\perp}$同样是$\mathcal{A}^*$不变的. 算子$\mathcal{A}$和$\mathcal{A}^*$在 $\left(V^{\lambda}\right)^{\perp}$ 上的限制显然是可以交换的, 从而是正规的. 对$n = \dim\ V$ 用归纳法, 我们可以认为$\mathcal{A}$ 在$\left(V^{\lambda}\right)^{\perp}$ 上是可对角化的. 按定义即可确信$V = V^{\lambda} \oplus \left(V^{\lambda}\right)^{\perp}$, 这就完成了证明. □

因为埃尔米特算子和酉算子都是正规的.所以按定理11可将它们对角化, 我们就很容易地得到了定理6和定理9中所说的谱性质.现在, 回顾一下第2章§3的定理1中的等方算子的完全正交组(投影).可以用下述方式描写对于正规算子的一般的谱定理.

定理12 对有限维空间V上的每个正规算子$\mathcal{A}$, 都有两两不同的特征值$\lambda_1, \cdots, \lambda_m$, $1 \leqslant m \leqslant n = \dim V$, 以及相互正交的投影算子$\mathcal{P}_1, \cdots, \mathcal{P}_m$, 使得

i) $\sum\limits_j \mathcal{P}_j = \mathcal{E}$;

ii) $\sum\limits_j \lambda_j \mathcal{P}_j = \mathcal{A}$ 是算子$\mathcal{A}$的谱分解式, $\lambda_j \in \mathrm{Spec}\,(\mathcal{A})$;

iii) 上述分解式是唯一的;

iv) 存在复多项式$f_1\,(t), \cdots, f_m\,(t)$具有性质

$$f_i\,(\lambda_j) = \delta_{ij}, \quad f_i\,(\mathcal{A}) = \mathcal{P}_i$$

(在自共轭算子情形, 所有的数λ_i和多项式$f_i\,(t)$都是实的).

证明 设$\lambda_1, \cdots, \lambda_m$是算子$\mathcal{A}$的所有的两两不同的特征值, $\mathcal{P}_i$是在V^{λ_i}上平行于$\sum\limits_{j \neq i} V^{\lambda_j}$的投影$(i = 1, 2, \cdots, m)$. 据定理11(同样看它的证明), 所有$\mathcal{P}_j$是相互正交的(也就是, $\mathcal{P}_i\mathcal{P}_j = \mathcal{P}_j\mathcal{P}_i = \delta_{ij}\mathcal{P}_i$), 且均不为$\mathcal{O}$. 其次, $V = \bigoplus\limits_i V^{\lambda_i}$, 故$\sum\limits_i \mathcal{P}_i = \mathcal{E}$, 这就是, $\mathcal{P}_1, \cdots, \mathcal{P}_m$为完全正交组的根据(断言i)).

对任意向量$\mathbf{v} \in V$, 我们有$\mathcal{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$, 其中$\mathbf{v}_j = \mathcal{P}_j \mathbf{v}$.这样一来$\mathcal{A}\mathbf{v} = \mathcal{A}\,(\mathcal{E}\mathbf{v}) = \mathcal{A}\left(\sum\limits_j \mathcal{P}_j \mathbf{v}\right) = \sum\limits_j \mathcal{A}\mathbf{v}_j = \sum\limits_j \lambda_j \mathbf{v}_j = \sum\limits_j \lambda_j\,(\mathcal{P}_j \mathbf{v}) = \left(\sum\limits_j \lambda_j \mathcal{P}_j\right)\mathbf{v}$, 而这就是基本断言ii).

至于有关算子$\mathcal{A}$的谱分解的唯一性断言iii), 我们可以这样来推导. 由$\mathcal{P}_i \neq \mathcal{O}$得到$\mathbf{x} \neq \mathbf{0}$, $\mathbf{x} \in \mathrm{Im}\ \mathcal{P}_i$的存在性. 按定义, $\mathcal{P}_i\ \mathbf{x} = \mathbf{x}$, 当$j \neq i$时, $\mathcal{P}_j\ \mathbf{x} = \mathbf{0}$. 所以,

$$\mathcal{A}\mathbf{x} = \sum_j \lambda_j \mathcal{P}_j \mathbf{x} = \lambda_i \mathcal{P}_i \mathbf{x} = \lambda_i \mathbf{x},$$

即, $\lambda_i \in \mathrm{Spec}\,(\mathcal{A})$.

反过来, 如果$\lambda \in \operatorname{Spec}(\mathcal{A})$且对某个$\mathbf{v} \in V, \mathbf{v} \neq \mathbf{0}$, 有$\mathcal{A}\mathbf{v} = \lambda\mathbf{v}$, 那么,

$$\mathcal{A}\mathbf{v} = \lambda\mathbf{v} = \lambda\sum_j \mathbf{v}_j, \quad \mathbf{v}_j = \mathcal{P}_j\mathbf{v},$$

而另一方面,

$$\mathcal{A}\mathbf{v} = \mathcal{A}\sum_j \mathbf{v}_j = \sum_j \mathcal{A}\mathbf{v}_j = \sum_j \lambda_j\mathbf{v}_j.$$

所以, $\sum_j (\lambda - \lambda_j)\mathbf{v}_j = \mathbf{0}$. 但是, 向量$\mathbf{v}_1, \cdots, \mathbf{v}_m$ 是相互正交的(这是投影算子$\mathcal{P}_1, \cdots, \mathcal{P}_m$ 相互正交的推论), 这意味着, 它们之中的那些个非零的向量必然是线性无关的. 可见, 对每个j都有$(\lambda - \lambda_j)\mathbf{v}_j = \mathbf{0}$, 而且如果$\mathbf{v}_i \neq \mathbf{0}$(这样的$i$一定能找到, 因为$\mathbf{v} \neq \mathbf{0}$), 就有$\lambda = \lambda_i \in \{\lambda_1, \cdots, \lambda_m\}$. 证明了唯一性.

断言iv)中的多项式$f_1(t), \cdots, f_m(t) \in \mathbb{C}(t)$可以具体构作成

$$f_i(t) = \prod_{j \neq i} \frac{t - \lambda_j}{\lambda_i - \lambda_j}.$$

显然, 如果$\mathcal{A}$是自共轭算子, 则$f_i(t) \in \mathbb{R}[t]$.

利用相互正交的投影$\mathcal{P}_j$组的定义和分解式ii), 我们有

$$\mathcal{A}^2 = \left(\sum_i \lambda_i\mathcal{P}_i\right)\left(\sum_j \lambda_j\mathcal{P}_j\right) = \sum_{i,j} \lambda_i\lambda_j\mathcal{P}_i\mathcal{P}_j = \sum_j \lambda_j^2\mathcal{P}_j,$$

$$\mathcal{A}^3 = \mathcal{A}\mathcal{A}^2 = \left(\sum_i \lambda_i\mathcal{P}_i\right)\left(\sum_j \lambda_j^2\mathcal{P}_j\right) = \sum_j \lambda_j^3\mathcal{P}_j,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathcal{A}^k = \sum_j \lambda_j^k\mathcal{P}_j$$

(当k=0时, 运用断言 i): $\mathcal{A}^0 = \sum_j \lambda_j^0\mathcal{P}_j = \sum_j \mathcal{P}_j = \mathcal{E}$). 这样一来, 对任意多项式$f(t)$都有

$$f(\mathcal{A}) = \sum_j f(\lambda_j)\mathcal{P}_j.$$

特别地,

$$f_i(\mathcal{A}) = \sum_j f_i(\lambda_j)\mathcal{P}_j = f_i(\lambda_i)\mathcal{P}_i = \mathcal{P}_i. \qquad \square$$

如同所有的线性算子一样, 正规算子$\mathcal{A}$可以写成$\mathcal{A} = \mathcal{B} + i\mathcal{C}$, $\mathcal{A}^* = \mathcal{B} - i\mathcal{C}$的样子, 其中$\mathcal{B}$和$\mathcal{C}$都是埃尔米特算子(见(6)).而它们本身同样可由$\mathcal{A}$和$\mathcal{A}^*$表达出来

$$\mathcal{B} = \frac{1}{2}(\mathcal{A} + \mathcal{A}^*), \quad \mathcal{C} = \frac{1}{2i}(\mathcal{A} - \mathcal{A}^*).$$

由$\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$得到$\mathcal{B}\mathcal{C} = \mathcal{C}\mathcal{B}$.反之, 由$\mathcal{B}$和$\mathcal{C}$的可以换位的性质可以导出$\mathcal{A}$和$\mathcal{A}^*$可以换位的性质

$$\mathcal{A}\mathcal{A}^* = \mathcal{B}^2 + \mathcal{C}^2 = \mathcal{A}^*\mathcal{A},$$

即$\mathcal{A}$的正规性质.

暂且把正规算子放下, 停在可交换位置的算子上, 或者, 如同前已说过的, 可交换的算子上.

引理6 设$\mathcal{A}$, $\mathcal{B}$是复空间V上两个可交换位置的算子.那么, $\mathcal{A}$和$\mathcal{B}$必有共同的特征向量.

证明 设$\lambda \in \mathrm{Spec}\,(\mathcal{A})$. 看子空间$V^\lambda = \{\mathbf{x} \in V | \mathcal{A}\mathbf{x} = \lambda\mathbf{x}\}$. 那么, $\mathcal{B}\, V^\lambda \subseteq V^\lambda$. 事实上, 利用条件$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, 我们得到蕴涵式

$$\mathbf{x} \in V^\lambda \Rightarrow \mathcal{A}\,(\mathcal{B}\mathbf{x}) = \mathcal{B}\,(\mathcal{A}\mathbf{x}) = \mathcal{B}\,(\lambda\mathbf{x}) = \lambda\,(\mathcal{B}\mathbf{x})\,,$$

也就是$\mathcal{B}\mathbf{x} \in V^\lambda$.

线性算子$\mathcal{B}$限制在V^λ上, 必有特征向量$\mathbf{y} \in V^\lambda : \mathcal{B}\mathbf{y} = \mu\mathbf{y}$, $\mu \in \mathrm{Spec}\,(\mathcal{B})$. 这样一来, $\mathcal{A}\mathbf{y} = \lambda\mathbf{y}$, $\mathcal{B}\mathbf{y} = \lambda\mathbf{y}$, 也就是说$\mathbf{y}$是它们的共同的特征向量. □

定理13 在n维埃尔米特空间V上, 两个埃尔米特算子$\mathcal{A}$, $\mathcal{B}$, 或者两个保距自同构$\mathcal{A}$, $\mathcal{B}$能在某个标准正交基底之下同时化成对角型, 当且仅当, 它们可以交换位置.

证明 假设$\mathcal{A}$和$\mathcal{B}$在一个共同的标准正交基底之下同时是对角的, 我们可以推出, 在此基底之下它们对应的矩阵A, B是可以交换位置的.但是, 因为在任意一个另外的基底之下算子的矩阵应该是$C^{-1}AC$, $C^{-1}BC$, 且

$$C^{-1}AC \cdot C^{-1}BC = C^{-1}ABC = C^{-1}BAC = C^{-1}BC \cdot C^{-1}AC,$$

所以, 算子本身是可以交换位置的.

反之, 设$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$. 那么, 据引理6, 算子$\mathcal{A}$, $\mathcal{B}$就有共同的特征向量$\mathbf{e}_1$. 不失一般性, 可以认为$\|\mathbf{e}_1\| = 1$.由于它们的埃尔米特性质(引理3)或者酉性质(引理5), 子空间$W = \langle\mathbf{e}_1\rangle^\perp$是$V$的$n-1$维的对于$\mathcal{A}$和$\mathcal{B}$都不变的子空间. $\mathcal{A}$和$\mathcal{B}$在W上的限制必然是可交换位置的埃尔米特算子(相应地, 酉算子).对维数用归纳法, 即可构造出来一个标准正交基底, 在这个基底之下, $\mathcal{A}$和$\mathcal{B}$同时表成对角形式. □

说明4 回忆一下, 据定理3, 埃尔米特算子$\mathcal{A}$, $\mathcal{B}$可以交换位置的性质等价于算子的$\mathcal{A}\mathcal{B}$的埃尔米特性质.

8. 正定算子 因为在埃尔米特空间V上的任意埃尔米特算子$\mathcal{A}$(欧几里得空间上的对称算子)都对应一个二次型$q\,(\mathbf{x}) = (\mathcal{A}\mathbf{x}|\mathbf{x})$, 于是, 那些可以应用于二次型上的概念, 如正定性, 半正定性等等都可以拿到算子$\mathcal{A}$上来.

定义6 埃尔米特(或线性对称的)算子$\mathcal{A}$称为是**正定的**, 如果对V中任意向量$\mathbf{x} \neq \mathbf{0}$, 都有$(\mathcal{A}\mathbf{x}|\mathbf{x}) > 0$.

定理6及其随后的说明指出, 对所有的正定算子$\mathcal{A}$都有V的标准正交基底, 在这个基底之下, 矩阵A具有对角形式

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \tag{14}$$

而且它的特征值$\lambda_1, \cdots, \lambda_n \in \mathbb{R}$都是正的. 反过来, 把任意一个形如(14)的矩阵A解释成一个埃尔米特(对称)算子$\mathcal{A}$相对于V的某个标准正交基底的矩阵, 我们可以得到结论, 条件$\lambda_1 > 0, \cdots, \lambda_n > 0$能保障$\mathcal{A}$的正定性.这种情形, 配备的符号$\mathcal{A} > 0$.

讨论算子$\mathcal{A}$的半正定性质(用$\mathcal{A} \geqslant 0$表示)也是有意义的, 这只要$\lambda_1 \geqslant 0, \cdots, \lambda_n \geqslant 0$并有某些$i$使$\lambda_i = 0$.两个埃尔米特(对称)算子$\mathcal{A}$, $\mathcal{B}$记成$\mathcal{A} \geqslant \mathcal{B}$, 只要$\mathcal{A} - \mathcal{B} \geqslant 0$.

按定义, 正定算子是非退化的(可逆的), 这可由柯西–布尼亚科夫斯基不等式

$$|(\mathcal{A}\mathbf{x}|\mathbf{x})| \leqslant \|\mathcal{A}\mathbf{x}\| \cdot \|\mathbf{x}\|$$

看出来. 反过来, 非退化条件及$\mathcal{A} \geqslant 0$可以限定$\mathcal{A}$的正定性.

命题1 *所有的正定算子$\mathcal{A}$均可以表达成某个另外的正定算子的平方: $\mathcal{A}=\mathcal{B}^2$, 其中平方根表达$\mathcal{B} := \sqrt{\mathcal{A}}$是唯一的.*

证明 把算子$\mathcal{A}$的矩阵化成形如(14)的对角形式并令$B = \mathrm{diag}\left(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}\right)$, 以及$\sqrt{\lambda_i} > 0$就足够了. 在给定的标准正交基底之下以$B$为其矩阵的线性算子$\mathcal{B}$必然是正定的. 关系式$A=B^2$在向另外的转化时仍然保持$C^{-1}AC = \left(C^{-1}BC\right)^2$. 于是有, $\mathcal{A} = \mathcal{B}^2$.

关于$\mathcal{B}$的唯一性的断言借助于关于谱分解的定理12, 证明起来是很方便的.也就是说, 如果$\mathcal{B}' > 0$且$(\mathcal{B}')^2 = \mathcal{A}$, 那么, 研究$\mathcal{B}'$的谱分解式$\mathcal{B}' = \sum\limits_j \mu_j \mathcal{P}'_j$, 我们就得到关系式

$$\sum_j \mu_j^2 \mathcal{P}'_j = (\mathcal{B}')^2 = \mathcal{A} = \sum_i \lambda_i \mathcal{P}_i.$$

所有的数$\mu_j > 0$两两不同, 它们的平方μ_j^2也是如此.算子$\mathcal{A}$的谱分解的唯一性提供我们一个结论, 集合$\{\mu_j^2\}$和$\{\lambda_i\}$是重合的, 也就是, 通过适当调整顺序, 应有等式$\mu_i^2 = \lambda_i$, $\mathcal{P}'_i = \mathcal{P}_i$, 从而$\mu_i = \sqrt{\lambda_i}$, 而且$\mathcal{B} = \mathcal{B}' = \sqrt{\mathcal{A}}$. □

命题2 *设$\mathcal{C}$是带有纯量乘积的空间上的任意一个非退化线性算子.那么, $\mathcal{A} = \mathcal{C}\mathcal{C}^*$(或$\mathcal{C}^*\mathcal{C}$)是个非退化的正定算子.*

证明 算子$\mathcal{C}\mathcal{C}^*$的埃尔米特性质(或对称性)已经验证过了: $(\mathcal{C}\mathcal{C}^*)^* = \mathcal{C}^{**}\mathcal{C}^* = \mathcal{C}\mathcal{C}^*$. 而$\mathcal{A} = \mathcal{C}\mathcal{C}^*$的非退化性是显然的: $\det\ \mathcal{C}^*\mathcal{C} = \det\ C^t\overline{C} = \det\ C\ \det\ \overline{C} = \det\ C\ \overline{\det\ C} = |\det\ C|^2 \neq 0$. 其次, $\mathbf{x} \neq \mathbf{0} \Rightarrow \mathcal{C}^*\mathbf{x} \neq \mathbf{0}$, 因此, 按共轭算子的定义, 我们有

$$(\mathcal{C}\mathcal{C}^*\mathbf{x}|\mathbf{x}) = (\mathcal{C}^*\mathbf{x}|\mathcal{C}^*\mathbf{x}) > 0, \quad \forall \mathbf{x} \neq \mathbf{0}.$$

这就意味着, $\mathcal{A} = \mathcal{C}\mathcal{C}^*$是个正定算子. 对于乘积$\mathcal{C}^*\mathcal{C}$也是一样. □

由建立在非退化算子情形上的命题1和命题2可以直接推导出

定理14 设V是带有纯量乘积$(*|*)$的空间. V上线性算子$\mathcal{A}$的下列条件是等价的:

i) $\mathcal{A} = \mathcal{B}^2, \mathcal{B}^* = \mathcal{B}$;

ii) $\mathcal{A} = \mathcal{C}\mathcal{C}^*$;

iii) $(\mathcal{A}\mathbf{x}|\mathbf{x}) \geqslant 0$.

在1维的复空间, 性质i)和性质ii)的每一个就是刻画一个非负实数: $z \geqslant 0$, 就意味着z可以写成$z = \lambda^2$, $\lambda \in \mathbb{R}$(性质i)的特殊情形), 以及$z = z'\overline{z'}$(性质ii)的特殊情形).

9. 极化分解 上面提到的复数与带有纯量乘积的空间上的线性算子之间的平行性质可以延伸得更远, 直到复数的三角公式: $z = |z|\, e^{i\varphi} = \sqrt{z\bar{z}}\, e^{i\varphi}$. 关于这一点, 可见证于

定理15 埃尔米特(或欧几里得)向量空间V上的每个非退化的线性算子$\mathcal{A}$都可表达成

$$\mathcal{A} = \mathcal{P}\mathcal{Q}, \tag{15}$$

的形式, 其中$\mathcal{P}$是个正定算子, 而$\mathcal{Q}$是个保距算子(酉算子或者正交算子). 分解式(15)是唯一的(称为算子$\mathcal{A}$的极化分解).

证明 按照命题1和命题2有$\mathcal{A}\mathcal{A}^* = \mathcal{P}^2$, 其中$\mathcal{P}$是个正定算子而且是唯一的平方根: $\mathcal{P} = \sqrt{\mathcal{A}\mathcal{A}^*}$. 当然, $\mathcal{P}$是可逆算子. 令$\mathcal{Q} = \mathcal{P}^{-1}\mathcal{A}$, 我们就得到表达式(15). 只需证明$\mathcal{Q}$是个保距算子.

事实上, 因为$\mathcal{P}^* = \mathcal{P}$而且$\mathcal{P}\,\mathcal{P}^{-1} = \mathcal{E} = \mathcal{E}^* = \left(\mathcal{P}^{-1}\right)^*\mathcal{P}^* \Rightarrow \left(\mathcal{P}^{-1}\right)^* = \left(\mathcal{P}^*\right)^{-1} = \mathcal{P}^{-1}$, 所以

$$\mathcal{Q} = \mathcal{P}^{-1}\mathcal{A}\mathcal{A}^*\left(\mathcal{P}^{-1}\right)^* = \mathcal{P}^{-1}\mathcal{P}^2\mathcal{P}^{-1} = \mathcal{E}.$$

现在, 如果有两个形如(15)的分解, $\mathcal{P}\mathcal{Q} = \mathcal{A} = \mathcal{P}_1\mathcal{Q}_1$, 那么, 我们有$\mathcal{Q}^*\mathcal{P} = \mathcal{Q}_1^*\mathcal{P}_1$. 所以. $\mathcal{P}\mathcal{Q}\cdot\mathcal{Q}^*\mathcal{P} = \mathcal{P}_1\mathcal{Q}_1\cdot\mathcal{Q}_1^*\mathcal{P}_1$, 从而$\mathcal{P}^2 = \mathcal{P}_1^2 \Rightarrow \mathcal{P} = \mathcal{P}_1$(平方根的唯一性), 而且, 由此可见, $\mathcal{Q} = \mathcal{Q}_1$, 即同样判明了极化分解式的唯一性. □

说明5 显然,

$$\mathcal{A} = \mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{Q}^{-1}\mathcal{P}\mathcal{Q},$$

由此我们看到

$$\mathcal{A} = \mathcal{Q}\mathcal{P}_1,$$

其中$\mathcal{Q}$是保距的, 而$\mathcal{P}_1 = \mathcal{Q}^*\mathcal{P}\mathcal{Q}$是个正定的线性算子.

当算子是退化的情形, 极化分解式(15)(不要求$\mathcal{Q}$有唯一性)也是对的.但是, 我们转向分解的其他性质.对复数z, 在其三角分解中, 因子的顺序可以不加区别: $|z|\, e^{i\varphi} = e^{i\varphi}\, |z|$.现在, 如果$\mathcal{A} = \mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P}$, 那么

$$\mathcal{A}\mathcal{A}^* = \mathcal{P}\mathcal{Q}\cdot\mathcal{Q}^*\mathcal{P}^* = \mathcal{P}^2 = \mathcal{P}\mathcal{Q}^*\mathcal{Q}\mathcal{P} = (\mathcal{Q}\mathcal{P})^*\,\mathcal{Q}\mathcal{P} = \mathcal{A}^*\mathcal{A},$$

这就意味着算子$\mathcal{A}$的正规性.

反过来,

$$\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A} \Rightarrow \mathcal{P}^2 = \mathcal{P}\mathcal{Q}\mathcal{Q}^*\mathcal{P} = \mathcal{Q}^*\mathcal{P}^*\mathcal{P}\mathcal{Q} = \mathcal{Q}^{-1}\mathcal{P}^2\mathcal{Q}.$$

但由$\mathcal{Q}$与$\mathcal{P}^2$的可换位性可推出$\mathcal{Q}$与$\mathcal{P}$的可换位性质, 因为$\mathcal{P} = \sqrt{\mathcal{P}^2}$是$\mathcal{P}^2$的多项式(可由谱分解定理得出来). 这样一来,

$$\mathcal{A} = \mathcal{P}\mathcal{Q}\text{是正规的} \Leftrightarrow \mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P}.$$

习 题

1. $n \times n$阶的酉矩阵A何时能表成(乘法)换位子$A = XYX^{-1}Y^{-1}$的形式, 其中X, Y也是酉矩阵?

2. 把**雅可比矩阵**理解为形如

$$J = \begin{pmatrix} a_1 & -b_1 & 0 & \cdots & 0 & 0 \\ -c_1 & a_2 & -b_2 & \cdots & 0 & 0 \\ 0 & -c_2 & a_3 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & a_{n-1} & -b_{n-1} \\ 0 & 0 & 0 & \cdots & -c_{n-1} & a_n \end{pmatrix}, \quad b_i c_i > 0, \quad 1 \leqslant i \leqslant n-1.$$

的实矩阵. 证明, Spec(J)永远是实的且都是单的.

3. 类似于定理13, 如果, 两个算子中一个是埃尔米特而第二个是保距的, 那么, 定理13还能保持正确吗?

4. 设$\mathcal{A}$, $\mathcal{B}$是域$\mathbb{K}$上向量空间V上的任意两个可交换的线性算子. 证明, 如果$\mathcal{A}$和$\mathcal{B}$均可对角化, 那么, 它们就可以同时被对角化, 也就是存在一个基底, 它由$\mathcal{A}$的特征向量组成, 同样, 它也是由$\mathcal{B}$的特征向量组成的.

5. 证明, 如果$\mathcal{A}$, $\mathcal{B}$都是正定线性算子且$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, 那么, $\mathcal{A}\mathcal{B}$也是正定的.

6. 如果${}^tA = -A$, 那么, 证明A^2必为对称的正定矩阵.特别地, 斜对称矩阵的非零特征值必为纯虚数.

7. 设$\mathcal{A}$, $\mathcal{B}$是埃尔米特(对称)算子, 其中之一, 比方说是$\mathcal{A}$, 是正定的.证明, Spec$(\mathcal{A}\mathcal{B})$是实的.

8. 设$q(\mathbf{x})$是带有纯量乘积$(*|*)$的欧几里得空间的一个二次型. 问, 在单位球面$(\mathbf{x}|\mathbf{x}) = 1$上的哪一个点上, 二次型$q$能达到极大和极小? 更一般地, 在单位球面的哪些点上二次型q取平稳值? 也就是在该点沿各个方向求导数均为零.证明, 下列论断是正确的.

二次型$q(\mathbf{x})$恰好在单位球面上那样一些点上取得平稳值, 这些点对应由型$q(\mathbf{x}) = (\mathcal{F}\mathbf{x}|\mathbf{x})$确定的对称算子$\mathcal{F}$的特征向量.

特别地, 型$q(\mathbf{x})$在单位球面上的极大值等于它的规范系数中的最大者, 而极小值就是最小的系数(二次型的极值).

9. 证明引理6的如下的延伸.

定理16 复的有限维向量空间上的任意一族可交换的线性算子必有共同的特征向量.

10. 如果

$$\mathfrak{G} = \{A_i \in M_n(\mathbb{C}) \,|\, A_iA_j = A_jA_i;\ i,j \in J\}$$

是任意一个可交换的(两两可换位置的)n阶矩阵的集合.那么，可以找到一个非退化矩阵C使得共轭集

$$\mathcal{C}^{-1}\mathfrak{G}\mathcal{C} = \left\{\mathcal{C}^{-1}A_jC|j \in J\right\}$$

是由可交换的上三角形矩阵组成的.

11. 按惯例，设E是n阶单位矩阵；E_{ij}，$1 \leqslant i,j \leqslant n$，是矩阵单位.验证，矩阵族

$$\mathfrak{G} = \left\{E_{ij}|1 \leqslant i \leqslant \left[n^2/2\right], \left[n^2/2\right] + 1 \leqslant j \leqslant n\right\} \cup \{E\}$$

含$\left[n^2/4\right] + 1$个元素而且是由线性无关的，交换的上三角形矩阵组成的.把$[p/q]$理解为分数p/q的整数部分.

12. **定理** [I.舒尔, 1905]. 在$M_n(\mathbb{C})$中可交换子代数的最大维数等于$\left[n^2/4\right] + 1$.

换句话说，需要证明，在$\mathbb{C}$上两两交换的线性无关的n阶矩阵最大的个数是$\left[n^2/4\right] + 1$.

事实上，$\mathbb{C}$可以用任意域$\mathfrak{K}$来代替.

13. 证明下列论断. 设$(V, (*|*))$是个偶数$n = 2m$维的欧几里得向量空间. 再设$f(\mathbf{x}, \mathbf{y})$是$V$上的非退化的斜对称型. 那么，必可将$V$分解成两个$m$维子空间的直和$V = V_1 \oplus V_2$，并能找到一个(对$(*|*)$)对称的非退化的线性算子$\mathcal{A} : V \to V$使得

$$f(\mathbf{x}, \mathbf{y}) = (\mathbf{x}_1|\mathcal{A}\mathbf{y}_2) - (\mathbf{x}_2|\mathcal{A}\mathbf{y}_1).$$

这里$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2, \mathbf{x}_i, \mathbf{y}_i \in V_i, i = 1, 2$.

§4 复化与实化

正如我们不止一次地相信了这样一种可能性，即要解决把线性算子$\mathcal{A} : V \to V$的矩阵按类别化成规范型的问题，取决于它的基础域是代数封闭的($\mathfrak{K} = \mathbb{C}$)还是不封闭的($\mathfrak{K} = \mathbb{R}$).特别地，它分别地对应了保距算子是酉算子和正交算子的情形.因为在复数情形代数图形(在损失某些几何直观的前提下)变得比较简单，所以常常把复数化算子(或者也说成是函子)用到实空间实算子，而借助于逆算子(实数化算子)再返回到原来开始时讨论的对象上去.我们仔细地把目光停在这里.

1. 复结构 设V是$\mathbb{R}$上的有限维(n维)向量空间.

定义1 如果给定一个线性算子$\mathcal{J} : V \to V$，其平方$\mathcal{J}^2 = -\mathcal{E}$，则说在$V$上定义了一个**复结构**.

例1 设$n=2m$, $\mathcal{J}$是个线性算子, 它对应矩阵

$$J=\begin{pmatrix} 0 & -1 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & -1 & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & -1 \\ & & & & & 1 & 0 \end{pmatrix} \tag{1}$$

(见第1章§4定理9的推论).显然, $\mathcal{J}$在V上定义了一个复结构, 因为$J^2=-E$.

词组"复结构"可以证实这样一种状况, 即对$(V,\ \mathcal{J})$可以变成一个$\mathbb{C}$上的向量空间$\widetilde{V}$, 令

$$(\alpha+i\beta)\mathbf{v}=\alpha\mathbf{v}+\beta\mathcal{J}\mathbf{v},\quad \alpha,\beta\in\mathbb{R},\quad \mathbf{v}\in V.$$

分配律公理

$$a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v},\quad (a+b)\mathbf{v}=a\mathbf{v}+b\mathbf{v},\quad a,b\in\mathbb{C},\quad \mathbf{u},\mathbf{v}\in V$$

得到满足, 因为$\mathcal{J}$是个线性算子. 其次, 由$\mathcal{J}^2=-\mathcal{E}$, 得到

$$\begin{aligned}(\alpha+i\beta)\left[(\gamma+i\delta)\mathbf{v}\right]&=(\alpha+i\beta)(\gamma\mathbf{v}+\delta\mathcal{J}\mathbf{v})=\alpha(\gamma\mathbf{v}+\delta\mathcal{J}\mathbf{v})+\beta\mathcal{J}(\gamma\mathbf{v}+\delta\mathcal{J}\mathbf{v})\\&=\alpha\gamma\mathbf{v}+\alpha\delta\mathcal{J}\mathbf{v}+\beta\gamma\mathcal{J}\mathbf{v}-\beta\delta\mathbf{v}=(\alpha\gamma-\beta\delta)\mathbf{v}+(\alpha\delta+\beta\gamma)\mathcal{J}\mathbf{v}\\&=\left[\alpha\gamma-\beta\delta+i(\alpha\delta+\beta\gamma)\right]\mathbf{v}=\left[(\alpha+i\beta)(\gamma+i\delta)\right]\mathbf{v}.\end{aligned}$$

所有的剩下的向量空间的公理也都被满足, 因为V和$\widetilde{V}$作为集合而言是一致的.

定义2 称$\widetilde{V}$是与实空间V相关的**复化向量空间**.

我们来证明, 例1不是偶然的.

命题1 有复结构的向量空间V在$\mathbb{R}$上必然是偶数维的, 而算子$\mathcal{J}$在某个基底之下的矩阵必形如(1).进一步,

$$\dim_{\mathbb{C}}\widetilde{V}=\frac{1}{2}\dim_{\mathbb{R}}V. \tag{2}$$

证明 假设我们已经找到了向量$\mathbf{e}_1,\cdots,\mathbf{e}_k$使得$2k$个向量$\mathbf{e}_1,\mathcal{J}\mathbf{e}_1,\cdots,\mathbf{e}_k,\mathcal{J}\mathbf{e}_k$原本是线性无关的了. 那么, 或者, 线性包络

$$V_k=\langle\mathbf{e}_1,\mathcal{J}\mathbf{e}_1,\cdots,\mathbf{e}_k,\mathcal{J}\mathbf{e}_k\rangle$$

已经与V重合, 或者, 如同早就证明过的那样, 还能找到$\mathbf{e}_{k+1}\notin V_k$.

立刻允许有

$$\mathcal{J}\mathbf{e}_{k+1}=\alpha\mathbf{e}_{k+1}+\mathbf{v}_k,\ \ \alpha\in\mathbb{R},\quad \mathbf{v}_k\in V_k.$$

于是, 我们可以把算子$\mathcal{J}$作用到这个等式的两端, 得

$$-\mathbf{e}_{k+1}=\alpha\mathcal{J}\mathbf{e}_{k+1}+\mathcal{J}\mathbf{v}_k$$

注意, 子空间V_k对于$\mathcal{J}$是不变的, 所以, $\mathcal{J}\mathbf{v}_k\in V_k$.把得到的第一个关系式乘以$\alpha$, 交换两边, 再减去第二个关系式, 就得到

$$(\alpha^2+1)\mathbf{e}_{k+1}=-\alpha\mathbf{v}_k-\mathcal{J}\mathbf{v}_k\in V_k.$$

但是, 这与$\mathbf{e}_{k+1}$的选法相矛盾, 故总有$\alpha^2+1=0$.

把这个往线性无关的向量集合V_k上添加线性无关的元素的过程一直继续下去, 最后, 我们一定会有某个m使得整个空间有

$$V=V_m=\langle\mathbf{e}_1,\mathcal{J}\mathbf{e}_1,\cdots,\mathbf{e}_m,\mathcal{J}\mathbf{e}_m\rangle\,.$$

这样一来, $\dim_{\mathbb{R}}V=2m$, 并且在基底$(\mathbf{e}_1,\mathcal{J}\mathbf{e}_1,\cdots,\mathbf{e}_m,\mathcal{J}\mathbf{e}_m)$之下算子$\mathcal{J}$的矩阵恰好有如(1)形式. 等式(2)则可以由向量$\mathbf{e}_k$和$\mathcal{J}\mathbf{e}_k$在$\mathbb{C}$上成比例$\mathcal{J}\mathbf{e}_k=i\mathbf{e}_k$这一事实直接得到. □

按存在性进行的推导重复了化斜对称型为规范型的过程(见第1章, §4).

2. 实化 现在设U是$\mathbb{C}$上的任意一个n维空间.

定义3 把U实数化并称为U的**实化空间**$U_{\mathbb{R}}$, 它作为集合, 作为加法群都与U重合, 但复数在它上面的乘法将被忘却, 而用实数去乘向量仍如原来U一样地实行.

我们把U简单化, 由n维空间

$$U=\langle\mathbf{e}_1,\cdots,\mathbf{e}_n\rangle_{\mathbb{C}}$$

得到$2n$维空间

$$U_{\mathbb{R}}=\langle\mathbf{e}_1,i\mathbf{e}_1,\cdots,\mathbf{e}_n,i\mathbf{e}_n\rangle_{\mathbb{R}}\,.$$

就把U实化了.

最初开始定义3的$i=\sqrt{-1}$在U上的乘法, 按给定的关系

$$\mathcal{J}\mathbf{e}_k=i\mathbf{e}_k,\mathcal{J}(i\mathbf{e}_k)=-\mathbf{e}_k,\quad 1\leqslant k\leqslant n. \tag{3}$$

就变成了在$U_{\mathbb{R}}$上的复结构, 就是线子$\mathcal{J}$.把在第1目中的看法用到对$(U_{\mathbb{R}},\mathcal{J})$上, 我们就得到了作为与出发点的空间$U_{\mathbb{R}}$相关的复化空间, 即

$$\widetilde{U_{\mathbb{R}}}=U.$$

现在, 我们引进如下的

定义4 线性算子$\mathcal{A}_{\mathbb{R}}:U_{\mathbb{R}}\to U_{\mathbb{R}}$, 它的作用与$\mathcal{A}$的作用是一样的, 被称为是线性算子$\mathcal{A}:V\to V$的**实化算子**. $\mathcal{A}$和$\mathcal{A}_{\mathbb{R}}$的差别可以归结为作用结果的解释上:

$$U_{\mathbb{R}}=\langle\mathbf{e}_1,\cdots,\mathbf{e}_n\rangle_{\mathbb{R}}+\langle i\mathbf{e}_1,\cdots,i\mathbf{e}_n\rangle_{\mathbb{R}}\,.$$

与这个分解式相对应的$\mathbb{C}$线性算子$\mathcal{A}: U \to U$, 我们记成, $\mathcal{A} = \mathcal{A}_1 + i\mathcal{A}_2$, 其中$\mathcal{A}_1$和$\mathcal{A}_2$ 是$\mathbb{R}$ 线性算子且在基底$\langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle$ 之下分别对应实的$n \times n$ 阶矩阵A_1 和A_2. 因为

$$\mathcal{A}(i\mathbf{e}_k) = i\mathcal{A}\mathbf{e}_k = i(\mathcal{A}_1\mathbf{e}_k + i\mathcal{A}_2\mathbf{e}_k) = -\mathcal{A}_2\mathbf{e}_k + i\mathcal{A}_1\mathbf{e}_k,$$

所以, 在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n; i\mathbf{e}_1, \cdots, i\mathbf{e}_n)$之下, 对$U_{\mathbb{R}}$实化算子$\mathcal{A}_{\mathbb{R}}$的矩阵是

$$A_{\mathbb{R}} = \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix}. \tag{4}$$

我们看到, 远非$U_{\mathbb{R}}$上的每个线性算子都可以作为U上的某个算子的实化算子.

设$\mathcal{L}(U)_{\mathbb{R}}$是所有的实化算子的集合, 而$\mathcal{L}(U_{\mathbb{R}})$是$U_{\mathbb{R}}$上所有的$\mathbb{R}$线性算子的向量空间, 由定义或者由实化算子的矩阵解释(4), 可以看出,

$$(\mathcal{A}+\mathcal{B})_{\mathbb{R}} = \mathcal{A}_{\mathbb{R}} + \mathcal{B}_{\mathbb{R}}, \quad (\mathcal{A}\mathcal{B})_{\mathbb{R}} = \mathcal{A}_{\mathbb{R}} \cdot \mathcal{B}_{\mathbb{R}}, \quad (\alpha\mathcal{A})_{\mathbb{R}} = \alpha\mathcal{A}_{\mathbb{R}}, \quad \alpha \in \mathbb{R}.$$

换句话说, $\mathcal{L}(U)_{\mathbb{R}}$是$\mathcal{L}(U_{\mathbb{R}})$的一个子代数. 显然,

$$\dim_{\mathbb{R}} \mathcal{L}(U)_{\mathbb{R}} = 2n^2 = \frac{1}{2}(2n)^2 = \frac{1}{2}\dim \mathcal{L}(U_{\mathbb{R}}).$$

在我们的基底之下, 线性算子$\mathcal{J}$(复结构)对应矩阵

$$J_0 = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}, \quad E = E_n. \tag{5}$$

按其意义, $\mathcal{A}_{\mathbb{R}} \cdot \mathcal{J} = \mathcal{J} \cdot \mathcal{A}_{\mathbb{R}}$, 它对应很容易验证的矩阵关系式$A_{\mathbb{R}} \cdot J_0 = J_0 \cdot A_{\mathbb{R}}$.再加上, 条件

$$\begin{pmatrix} A_1 & A_3 \\ A_2 & A_4 \end{pmatrix} \cdot \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix} = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix} \cdot \begin{pmatrix} A_1 & A_3 \\ A_2 & A_4 \end{pmatrix},$$

可以改写成(经分块矩阵的互换)形如

$$\begin{pmatrix} A_3 & -A_1 \\ A_4 & -A_2 \end{pmatrix} = \begin{pmatrix} -A_2 & -A_4 \\ A_1 & A_3 \end{pmatrix},$$

我们看到, $A_3 = -A_2$, $A_4 = A_1$. 也就是说, 所有的与J_0 可交换的$\mathbb{R}$上的$2n \times 2n$阶的矩阵必然形如(4).

于是, 有下面命题成立.

命题2 (对复结构$\mathcal{J}$)所有实化算子的子代数$\mathcal{L}(U)_{\mathbb{R}} \subset \mathcal{L}(U_{\mathbb{R}})$刚好由所有可与$\mathcal{J}$交换位置的算子组成.

更为有趣的是下面的问题. 设V是个实的维数为偶数的向量空间(比方说, $\dim_{\mathbb{R}} V = 2n$)且$\mathcal{A}: V \to V$是个线性算子. 在V上, 什么时候能存在一个复结

构$\mathcal{J}$能与$\mathcal{A}$相适当, 也就是, 使$\mathcal{A}=\mathcal{B}_{\mathbb{R}}$, 其中$\mathcal{B}: V \to V$是复的$n$维空间$U$上的线性算子? 我们未能触及部分内容精彩的情形.

定理1 设$V=\mathbb{R}^2$, 且$\mathcal{A}: V\to V$是个没有特征向量的$\mathbb{R}$线性算子. 那么, 在V上可以定义一个与$\mathcal{A}$相适的复结构(详细情况进入证明就可以知道).

证明 按条件$\mathcal{A}$有两个复共轭的特征根$\lambda, \bar{\lambda}$. 设$\lambda=\lambda_1+i\lambda_2$, $\lambda_1,\lambda_2\in\mathbb{R}$, 且$\lambda_2\neq 0$. 根据哈密顿–凯莱定理, $\mathcal{A}^2-(\mathrm{tr}\ \mathcal{A})\mathcal{A}+(\det\ \mathcal{A})\mathcal{E}=\mathcal{O}$, 也就是

$$\mathcal{A}^2-2\lambda_1\mathcal{A}+(\lambda_1^2+\lambda_2^2)\mathcal{E}=\mathcal{O}. \tag{6}$$

我们定义一个算子$\mathcal{J}$, 令

$$\mathcal{J}=\lambda_2^{-1}(\mathcal{A}-\lambda_1\mathcal{E}),$$

或者, 等价地,

$$\mathcal{A}=\lambda_1\mathcal{E}+\lambda_2\mathcal{J}.$$

代换等式(6)的$\mathcal{A}$, 我们得到

$$(\lambda_1^2\mathcal{E}+2\lambda_1\lambda_2\mathcal{J}+\lambda_2^2\mathcal{J}^2)-2\lambda_1(\lambda_1\mathcal{E}+\lambda_2\mathcal{J})+(\lambda_1^2+\lambda_2^2)\mathcal{E}=\mathcal{O},$$

进而推出

$$\mathcal{J}^2=-\mathcal{E}.$$

按照第2目的一般性推断, 在V上就定义了复的直线结构$\mathbb{C}^1$.因为算子$\mathcal{A}$可以与$\mathcal{J}$交换位置, 故$\mathcal{A}=\mathcal{B}_{\mathbb{R}}$, 其中$\mathcal{B}:\mathbb{C}^1\to\mathbb{C}^1$是一个用某个复数相乘的线性算子, 这个数, 显然, 就是λ. □

现在, 我们证明

命题3 $\det\ \mathcal{A}_{\mathbb{R}}=|\det\ \mathcal{A}|^2$.

证明 我们直接给出基于元素变换的计算出来的成果, 而不过分拘泥于实化的细节. 因为$\det\ \bar{A}=\overline{\det\ A}$, 其中小横杠代表取共轭复数, 所以有关系(4), 用记号$A=A_1+iA_2$, 那么

$$\begin{aligned}\det\ \mathcal{A}_{\mathbb{R}}&=\det\ A_{\mathbb{R}}=\det\begin{pmatrix}A_1 & -A_2\\ A_2 & A_1\end{pmatrix}\\ &=\det\begin{pmatrix}A_1+iA_2 & -A_2+iA_1\\ A_2 & A_1\end{pmatrix}=\det\begin{pmatrix}A_1+iA_2 & 0\\ A_2 & A_1-iA_2\end{pmatrix}\\ &=\det\begin{pmatrix}A & 0\\ A_2 & \bar{A}\end{pmatrix}=\det A\cdot\det\bar{A}=|\det A|^2.\end{aligned}$$ □

3. 复化 设V是$\mathbb{R}$上任意一个n维向量空间. 可以直接验证: 在外直和$V\oplus V$ (即向量空间的对$(V\ ,\ V)$)上带有运算

$$\alpha(\mathbf{u},\mathbf{v})+\alpha'(\mathbf{u}',\mathbf{v}')=(\alpha\mathbf{u}+\alpha'\mathbf{u}',\alpha\mathbf{v}+\alpha'\mathbf{v}'),\quad \alpha,\alpha'\in\mathbb{R},$$

且用对应

$$\mathcal{J} : (\mathbf{u}, \mathbf{v}) \mapsto (-\mathbf{v}, \mathbf{u})$$

定义了线性算子, 它给出了$V \oplus V$上的一个复结构.称这个复结构为**规范的**.

定义5 与$V \oplus V$相关联的复向量空间$\widetilde{V \oplus V}$被称为是V的**复化空间**(或者**复的包络**).

对此, 我们引入专门表示

$$V^{\mathbb{C}} := \widetilde{V \oplus V}.$$

如果把$\mathbb{C}$看成是$\mathbb{R}$上的2维向量空间.那么

$$V^{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C} \ ,$$

这是向量空间的张量乘积中的一种特殊情形, 它在数学中有广泛的应用(我们可以回顾一下第1章§4, 而更详细的研究将在第6章§2中进行). 因为$\dim_{\mathbb{R}}(V \oplus V) = 2n$, 所以等式(2)对应的是

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V.$$

按定义, $i(\mathbf{u}, \mathbf{v}) = \mathcal{J}(\mathbf{u}, \mathbf{v}) = -(\mathbf{v}, \mathbf{u})$, 所以, $(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \mathbf{0}) + i(\mathbf{v}, \mathbf{0})$.从而可将对$(\mathbf{u}, \mathbf{v})$很自然地表示成$\mathbf{u} + i\mathbf{v}$, 于是,

$$(\mathbf{u} + i\mathbf{v}) + (\mathbf{u}' + i\mathbf{v}') = (\mathbf{u} + \mathbf{u}') + i(\mathbf{v} + \mathbf{v}').$$

其次,

$$(\alpha + i\beta)(\mathbf{u} + i\mathbf{v}) = (\alpha\mathbf{u} - \beta\mathbf{v}) + (\alpha\mathbf{v} + \beta\mathbf{u}),$$

这是因为,

$$(\alpha\mathcal{E} + \beta\mathcal{J})(\mathbf{u}, \mathbf{v}) = \alpha(\mathbf{u}, \mathbf{v}) + \beta(-\mathbf{v}, \mathbf{u}) = (\alpha\mathbf{u} - \beta\mathbf{v}, \alpha\mathbf{v} + \beta\mathbf{u}).$$

要记住这些规则并不困难, 因为它们刚好对应作用在复数上的那些规则. 向量$\mathbf{u} + i\mathbf{0}$就简单地记成$\mathbf{u}$, 从而实空间V可以看成是$V^{\mathbb{C}}$的一个子空间.

定义6 对于$\mathbb{R}$线性算子$\mathcal{A} : V \to V$, 我们把$\mathbb{C}$线性算子$\mathcal{A}^{\mathbb{C}} : V \to V$,

$$\mathcal{A}^{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = \mathcal{A}\mathbf{u} + i\mathcal{A}\mathbf{v},$$

称为$\mathcal{A}$的**复化算子**.

我们有蕴涵式

$$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle_{\mathbb{R}} \Rightarrow V^{\mathbb{C}} = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle_{\mathbb{C}} .$$

由此可见, 算子$\mathcal{A}$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的矩阵$A$同时也就是算子$\mathcal{A}^{\mathbb{C}}$在同样的基底之下的矩阵$A^{\mathbb{C}}$. 即

$$A^{\mathbb{C}} = A.$$

特别地, $\det \mathcal{A}^{\mathbb{C}} = \det \mathcal{A}$且$\operatorname{tr} \mathcal{A}^{\mathbb{C}} = \operatorname{tr} \mathcal{A}$. 因为,

$$\begin{aligned}(\mathcal{A}+\mathcal{B})^{\mathbb{C}}(\mathbf{u}+i\mathbf{v}) &= (\mathcal{A}+\mathcal{B})\mathbf{u}+i(\mathcal{A}+\mathcal{B})\mathbf{v}\\ &= (\mathcal{A}\mathbf{u}+\mathcal{B}\mathbf{u})+i(\mathcal{A}\mathbf{v}+\mathcal{B}\mathbf{v}) = (\mathcal{A}\mathbf{u}+i\mathcal{A}\mathbf{v})+(\mathcal{B}\mathbf{u}+i\mathcal{B}\mathbf{v})\\ &= \mathcal{A}^{\mathbb{C}}(\mathbf{u}+i\mathbf{v})+\mathcal{B}^{\mathbb{C}}(\mathbf{u}+i\mathbf{v}) = (\mathcal{A}^{\mathbb{C}}+\mathcal{B}^{\mathbb{C}})(\mathbf{u}+i\mathbf{v}),\end{aligned}$$

所以, $(\mathcal{A}+\mathcal{B})^{\mathbb{C}} = \mathcal{A}^{\mathbb{C}}+\mathcal{B}^{\mathbb{C}}$.类似地可以验证, $(\mathcal{A}\mathcal{B})^{\mathbb{C}} = \mathcal{A}^{\mathbb{C}}\mathcal{B}^{\mathbb{C}}$.

类似于$\mathcal{A}^{\mathbb{C}}$继承了在V上的线性性质, 更一般地, 可以定义出$V^{\mathbb{C}}$的半线性型.例如, 如果f是向量空间V上的双线性型, 那么, 可令

$$f^{\mathbb{C}}(\mathbf{x}+i\mathbf{y},\mathbf{u}+i\mathbf{v}) = f(\mathbf{x},\mathbf{u})-f(\mathbf{y},\mathbf{v})+i(f(\mathbf{x},\mathbf{v})+f(\mathbf{y},\mathbf{u})).$$

作为一个习题, 请验证, 由f的斜对称性可以推出$f^{\mathbb{C}}$的斜对称性.

现在设V是个带有纯量乘积$(*|*)$的实向量空间, 那么, 也可以在$V^{\mathbb{C}}$上定义纯量乘积

$$(\mathbf{x}+i\mathbf{y}|\mathbf{u}+i\mathbf{v})^{\mathbb{C}} := (\mathbf{x}|\mathbf{u})+(\mathbf{y}|\mathbf{v})-i((\mathbf{x}|\mathbf{v})-(\mathbf{u}|\mathbf{y})).$$

与此同时, 如果对$(V,(*|*))$是个欧几里得空间, 那么, 对$(V^{\mathbb{C}},(*|*)^{\mathbb{C}})$就是个埃尔米特向量空间. 特别地, 在$V^{\mathbb{C}}$上的模$\|*\|^{\mathbb{C}}$将用等式

$$(\|\mathbf{x}+i\mathbf{y}\|^{\mathbb{C}})^2 = \|\mathbf{x}\|^2+\|\mathbf{y}\|^2$$

给出.

返回到一般情形, 如果$\mathcal{A}$是V上的一个线性算子, 而$\mathbf{a}+i\mathbf{b}$是$V^{\mathbb{C}}$上线性算子$\mathcal{A}^{\mathbb{C}}$的属于特征值$\alpha+i\beta$的一个特征向量($\mathbf{a},\mathbf{b}\in V,\alpha,\beta\in\mathbb{R}$). 那么, 与定义相对应, $\mathcal{A}\mathbf{a}+i\mathcal{A}\mathbf{b} = \mathcal{A}^{\mathbb{C}}(\mathbf{a}+i\mathbf{b}) = (\alpha+i\beta)(\mathbf{a}+i\mathbf{b}) = (\alpha\mathbf{a}-\beta\mathbf{b})+i(\beta\mathbf{a}+\alpha\mathbf{b})$,也就是,

$$\mathcal{A}\mathbf{a} = \alpha\mathbf{a}-\beta\mathbf{b},\quad \mathcal{A}\mathbf{b} = \beta\mathbf{a}+\alpha\mathbf{b}.$$

这样一来, $\langle\mathbf{a},\mathbf{b}\rangle_{\mathbb{R}}$就是一个对于$\mathcal{A}$不变的2维的子空间. 因为$\mathcal{A}^{\mathbb{C}}$总是至少要有一个特征向量, 所以, 我们再次证明了第2章§3的定理7.

其次, 要注意, $\mathbb{C}$上的每一个n维空间U必然同构于一个复包络$V^{\mathbb{C}}$, 它是由在$\mathbb{R}$上挑选出来的向量空间V导出来的. 这只要在U中确定某个基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$, V作为所有形如$\sum_j \alpha_j\,\mathbf{e}_j$, $\alpha_i\in\mathbb{R}$ 的向量的集合.

$$U = \langle\mathbf{e}_1,\cdots,\mathbf{e}_n\rangle_{\mathbb{C}} = (\langle\mathbf{e}_1,\cdots,\mathbf{e}_n\rangle_{\mathbb{R}})^{\mathbb{C}}.$$

4. 复化—实化—复化 把一个n维的实空间V复化后再实化就得到一个$2n$维的实的空间, 我们引入记号

$$W = (V^{\mathbb{C}})_{\mathbb{R}}.$$

容易相信,

$$W = V\oplus iV, \tag{7}$$

在这里, 说V是**实平面**而iV是**虚平面**. 对照(3), 我们在W上定义了算子$\mathcal{J}=(i\mathcal{E})_{\mathbb{R}}$, 它是在$V^{\mathbb{C}}$中用$i$相乘所得到的算子$i\mathcal{E}$的实化. J_0是它的矩阵(见(5)). 算子$\mathcal{J}$把实平面和虚平面交换位置.

最简单的情形是$V=\mathbb{R}, V^{\mathbb{C}}=\mathbb{C}^1$; 进而$W=(\mathbb{C}^1)_{\mathbb{R}}=\mathbb{R}^2$的情形. 在复的直线上定义了复数取共轭数的算子

$$\alpha+i\beta\mapsto\overline{\alpha+i\beta}=\alpha-i\beta.$$

对于一般情形, 在空间(7)上, 有类似的线性算子

$$\mathcal{S}:\mathbf{u}+i\mathbf{v}\mapsto\overline{\mathbf{u}+i\mathbf{v}}=\mathbf{u}-i\mathbf{v}$$

它对应的矩阵是

$$S=\begin{pmatrix}E_n & 0\\ 0 & -E_n\end{pmatrix}.$$

拓展这种情形, 研究任意一个$\mathbb{C}$线性算子$\mathcal{A}:V^{\mathbb{C}}\to V^{\mathbb{C}}$ (这个想法甚至可以用到线性映射$\mathcal{A}:V_1^{\mathbb{C}}\to V_2^{\mathbb{C}}, V_1\neq V_2$的情形). 称算子$\overline{\mathcal{A}}:V^{\mathbb{C}}\to V^{\mathbb{C}}$.

$$\overline{\mathcal{A}}\cdot\mathbf{u}+i\mathbf{v}=\overline{\mathcal{A}(\mathbf{u}+i\mathbf{v})},$$

是$\mathcal{A}$的**复共轭算子**. 此时,

$$(\overline{\mathcal{A}})_{\mathbb{R}}=\mathcal{S}\cdot\mathcal{A}_{\mathbb{R}}\cdot\mathcal{S}.$$

算子$\overline{\mathcal{A}}$在$\mathbb{C}$上的线性性质可以由$\mathcal{A}$的线性性质和复数取共轭的线性性直接得到. 记线性算子$\mathcal{A}$在空间V的基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下的矩阵是$A=A_1+iA_2$, 而算子$\overline{\mathcal{A}}$的矩阵是$\overline{A}=A_1-iA_2$, 其中$A_1$和$A_2$都是实矩阵(见第2目的推理). 由此得知, $\mathcal{A}=\overline{\mathcal{A}}$的充要条件是可以把它写成$\mathcal{A}=\mathcal{B}^{\mathbb{C}}$ (某个$V\to V$的实算子的复化).利用复数取共轭的算子的概念, 我们对有形如(4)矩阵的任意的实算子$\mathcal{A}_{\mathbb{R}}$可以写出

$$\operatorname{tr}\mathcal{A}_{\mathbb{R}}=2\operatorname{tr}A_1=\operatorname{tr}(\mathcal{A}+\overline{\mathcal{A}})=\operatorname{tr}\mathcal{A}+\operatorname{tr}\overline{\mathcal{A}}.$$

设U是任意一个复空间. 在第1章§4的第1目中, 我们已经讨论了空间$\overline{U}$的例子, 并称它是**复共轭向量空间**, 它与U的区别仅仅在于纯量乘积: $\lambda\otimes\mathbf{x}=\overline{\lambda}\mathbf{x}$.

类似地, 如果$(V,\mathcal{J})$是实空间附加上了复结构, 那么, 线性算子$-\mathcal{J}$也定义了一个复结构, 称它是**带初始端的复结构**. 其次, 如果$\widetilde{V}$是$(V,\mathcal{J})$对应的复空间, 那么, $\overline{\widetilde{V}}$就是$(V,-\mathcal{J})$对应的复空间.

现在, 我们先把实化函子, 继而再把复化函子用到复向量空间V上, 就建立了规范的$\mathbb{C}$线性同构

$$f:(V_{\mathbb{R}})^{\mathbb{C}}\to V\oplus\overline{V}.$$

为此, 还需要注意, 在$(V_{\mathbb{R}})^{\mathbb{C}}$上有两个$\mathbb{R}$线性算子: 规范的复结构算子$\mathcal{J}(\mathbf{x},\mathbf{y})=(-\mathbf{y},\mathbf{x})$和乘以$i=\sqrt{-1}$的算子, 它对应$V$上带初始端的复结构: $i(\mathbf{x},\mathbf{y})=(i\mathbf{x},i\mathbf{y})$. 因为$\mathcal{J}$与$i$可以交换, 所以, 在这个结构上它是$\mathbb{C}$线性的. 因为$\mathcal{J}^2=-\mathcal{E}$, 所以它的特征值是$\pm i$. 我们引进相应于这些特征值的两个特征子空间的规范表示:

$$V^{1,0}=\left\{(\mathbf{x},\mathbf{y})\in(V_{\mathbb{R}})^{\mathbb{C}}|\mathcal{J}(\mathbf{x},\mathbf{y})=i(\mathbf{x},\mathbf{y})\right\}.$$
$$V^{0,1}=\left\{(\mathbf{x},\mathbf{y})\in(V_{\mathbb{R}})^{\mathbb{C}}|\mathcal{J}(\mathbf{x},\mathbf{y})=-i(\mathbf{x},\mathbf{y})\right\}.$$

两个集合$V^{1,0}$和$V^{0,1}$, 都是$(V_{\mathbb{R}})^{\mathbb{C}}$的复的子空间: 显然, 它们对加法和用实数乘都是封闭的. 而相对于用$\mathcal{J}$乘的封闭性质可由$\mathcal{J}$和i的交换性得到.于是, 我们可以说$V=V^{1,0}\oplus V^{0,1}$, 还可以说, $V^{1,0}$自然同构于V, 而$V^{0,1}$自然同构于$\overline{V}$.

由定义立刻可以得到, $V^{1,0}$是由向量$(\mathbf{x},-i\mathbf{x})$组成的, 而$V^{0,1}$是由形如$(\mathbf{y},i\mathbf{y})$的向量组成的. 对于给定的$\mathbf{u}$, $\mathbf{v}\in V$, 方程$(\mathbf{u},\mathbf{v})=(\mathbf{x},-i\mathbf{x})+(\mathbf{y},i\mathbf{y})$有唯一解, $\mathbf{x}=(\mathbf{u}+i\mathbf{v})/2,\mathbf{y}=(\mathbf{u}-i\mathbf{v})/2$. 进而$V=V^{1,0}\oplus V^{0,1}$. 映射$\mathbf{x}\mapsto(\mathbf{x},-i\mathbf{x})$和$\mathbf{x}\mapsto(\mathbf{x},i\mathbf{x})$分别是$V$到$V^{1,0}$和$V$到$V^{0,1}$的$\mathbb{R}$线性同构. 此外, 由定义, 它们在$V$和$\overline{V}$上分别与$i$作用可以交换位置, 在$V^{1,0}$和$V^{0,1}$上可以与$\mathcal{J}$作用交换位置. 这就完成了我们的构建.

习　　题

1. 证实, 欧几里得空间V上没有特征向量的正交算子$\mathcal{A}$(它只能出现在$\dim V=2m$的情形), 一定是一个与V相关联的m维复向量空间U上的某个酉算子$\mathcal{B}:U\to U$的实化算子.在这种联系中, 要注意, 把酉空间实化就导出一个维数翻一番的欧几里得空间.

2. 在复空间U中选择一个基底, 在这个基底之下, 算子$\mathcal{A}:V\to V$的矩阵是个上三角形, 对角线上的元素是$\lambda_1,\cdots,\lambda_n$.证明命题3的公式.

3. 设U是$\mathbb{C}$上的向量空间.复化空间$V^{\mathbb{C}}$同构于什么?

4. 设$(V,(*|*))$是个欧几里得空间, $V^{\mathbb{C}}$是它的复化空间, $\mathcal{A}$是$V^{\mathbb{C}}$上的线性算子, 它的定义法则是, 对任意$\mathbf{u},\mathbf{v}\in V$, $\mathcal{A}(\mathbf{u}+i\mathbf{v})=\mathbf{u}-i\mathbf{v}$. 问, $\mathcal{A}$是实化空间$\left(V^{\mathbb{C}}\right)_{\mathbb{R}}$上的线性算子吗? 如果答案是肯定的, 再问, $\mathcal{A}$是对称的吗? 是正交的吗? 是等方的吗?

§5 正交多项式

1. 逼近问题 在数学和物理学里极其不同的问题中都会遇到在某个函数类中任取其中之一然后按给定的函数系分解的问题.不热衷于分析精确程度, 那通常是在分析学教程中研究的, 我们只限于纯粹地讨论这个问题的代数方面, 讨论将会顺便触及到某些新的线性代数和几何问题.

我们假设空间$C_2(a,b)$包含了所有的实变量t的在线段$a\leqslant t\leqslant b$上的连续函数(或者在以无穷远点为终点的区间上的连续函数), 它带有纯量乘积

$$(f|g)=\int_a^b f(t)\overline{g(t)}dt.$$

当我们遇到必须要研究复函数的情形, 小横杠表示取共轭数.在$C_2(a,b)$中将重点选出一些光滑函数的子类, 例如, 二次连续可微函数构成的子类.

用通常的办法引进函数f的模: $\|f\|=\sqrt{f|f}$.用距离

$$d(f,g)=\|f-g\|$$

就把空间变成可度量了. 前面说过的一般问题的表述可基于对标准正交函数系$\varphi_1(t),\varphi_2(t),\cdots$

$$(\varphi_i|\varphi_j)=\delta_{ij}$$

的研究以及对受"可逼近的"函数$f\in C_2(a,b)$支配的, 用系数α_i做出的线性组合$\sum\limits_i\alpha_i\varphi_i(t)$的研究.逼近问题仅仅在这样一种情况下才有意义, 那就是函数$\varphi_i(t)$足够好, 它是无穷次可微的, 甚至是解析的.

如果函数$f(t)$出现在物理学或力学的某种周期过程中, 那么, 自然地借助初等周期函数sin nt, cos nt, $n=0,1,\cdots$ 建立函数系$\{\varphi_n(t)\}$. 一般情形, 通常的多项式空间$\mathbb{R}[t]$ 可以充当标准正交系的一个很好的来源. 函数系$\{\varphi_n(t)\}$的建立可以简单地归结为逐次地运用我们已经熟悉了的格拉姆–施密特正交化过程. 我们很快就要研究这两个例子(sin, cos和$\mathbb{R}[t]$), 而现在停下来更准确地阐述一下对函数$f(t)$的逼近问题(也说是近似方法).

2. 最小二乘法 设给定了一个标准正交的函数系$\{\varphi_n(t)\}$. 如果f是$C_2(a,b)$中的任意一个函数, 那么, 数

$$c_n=(f|\varphi_n),n=1,2,\cdots,$$

被称为是函数f对于$\{\varphi_n(t)\}$的**傅里叶系数**. 因为

$$\begin{aligned}0\leqslant\left\|f-\sum_{j=1}^{n}c_j\varphi_j\right\|^2&=\left(f-\sum_{j=1}^{n}c_j\varphi_j|f-\sum_{s=1}^{n}c_s\varphi_s\right)\\&=\|f\|^2-\sum_j c_j(\varphi_j|f)-\sum_s\overline{c_s}(f|\varphi_s)+\sum_{j,s}c_j\overline{c_s}(\varphi_j|\varphi_s)\\&=\|f\|^2-\sum_j c_j\overline{(f|\varphi_j)}-\sum_s\overline{c_s}c_s+\sum_{j,s}c_j\overline{c_s}\delta_{js}\\&=\|f\|^2-\sum_j c_j\overline{c_j}-\sum_s\overline{c_s}c_s+\sum_j c_j\overline{c_j}=\|f\|^2-\sum_{j=1}^{n}|c_j|^2,\end{aligned}$$

所以, 永远有

$$\sum_{j=1}^{n}|c_j|^2\leqslant\|f\|^2.$$

不等式的右端与数字n没有关系, 因而, 事实上, 就是满足不等式

$$\sum_{k=1}^{\infty}|c_k|^2\leqslant\|f\|^2,\tag{1}$$

或者, 同样地,

$$\sum_{k=1}^{\infty} (f|\varphi_k)^2 \leqslant \|f\|^2. \tag{1'}$$

这个对任意标准正交系都成立的不等式(1)被称为**贝塞尔不等式**. 它证明了具有非负项的级数$\sum_{k\geqslant 1} |c_k|^2$的收敛性.

在最小二乘法意义下, 对给定的函数$f(t)$用一个具有不变系数d_k和预先给定的个数m的线性组合$\sum_{k=1}^{m} d_k\varphi_k(t)$去逼近, 意味着, 选取系数$d_k$使得平均平方差$\left\| f - \sum_{k=1}^{m} d_k\ \varphi_k(t) \right\|^2$最小(按另外的术语, 得到最小的平均的误差平方). 这个问题的几何意义是相当明显的. 对于(无穷维的)向量空间$V = C_2(a,b)$中的向量f, 我们在线性包络$U = \langle\varphi_1, \varphi_2, \cdots, \varphi_m\rangle$上构造一个向量$\varphi$, 它到$f$的距离$\|f-\varphi\|$是一个极小值. 这也同样被称为是**点到子空间的最短距离问题**, 或者, 同样地称为**垂直线问题**. 晚些时候, 我们还会碰到这个问题, 那是在真正点空间里面, 而不是在向量空间中. 我们总有直和分解

$$V = U \oplus U^{\perp},$$

因为$f = f_0 + f_1$, 其中f_0是f在U上的投影, 而f_1是"由终点向量f"到U上的垂直线(不再有任何意义的表达), 或者, 同样地, 是f在$U^{\perp}$上的投影. 现在, 如果$\varphi \in U$, $\varphi \neq f_0$, 那么,

$$\|f-\varphi\| > \|f - f_0\|. \tag{2}$$

事实上, $f_0 - \varphi \in U$并进而有$f_0 - \varphi$正交于$f_1 = f - f_0$. 按照毕达哥拉斯定理

$$\|f-\varphi\|^2 = \|f - f_0 + f_0 - \varphi\|^2 = \|f - f_0\|^2 + \|f_0 - \varphi\|^2,$$

从而得到不等式(2).

关于垂直线的实际问题可以转化为寻找向量f在U上的投影f_0的问题. 把f_0记成

$$f_0 = x_1\varphi_1 + \cdots + x_m\varphi_m,$$

形式. 由条件

$$(f - f_0|\varphi_j) = 0, \quad j = 1, 2, \cdots, m,$$

再由可以表示成m个未知数的线性方程组

$$x_1(\varphi_1|\varphi_j) + x_2(\varphi_2|\varphi_j) + \cdots + x_m(\varphi_m|\varphi_j) = (f|\varphi_j), \quad 1 \leqslant j \leqslant m \tag{3}$$

形式的条件, 我们就找出了未知的系数x_i.

条件(3)就是在函数系$\{\varphi_1,\cdots,\varphi_m\}$没有标准正交化的时候也能成立. 如果它已经标准正交化, 那么, $c_j=(f|\varphi_j)$就是傅里叶系数, 而且系(3)立刻给出

$$x_j=(f|\varphi_j)=c_j,\quad 1\leqslant j\leqslant m.$$

回到我们的近似方法问题, 可以得到这样的结论, 平均方差条件$\|f-\sum d_j\varphi_j\|^2$, 当$d_j=c_j$时将会取得极小值.顺便说一句, 这也可以直接看出

$$\left\|f-\sum_{j=1}^{m}d_j\varphi_j\right\|^2=\|f\|^2-\sum_{j=1}^{m}|c_j|^2+\sum_{j=1}^{m}|d_j-c_j|^2.$$

如果, 当m增大时, 对任意函数$f\in C_2(a,b)$, 差模$\left\|f-\sum\limits_{j=1}^{m}(f|\varphi_j)\varphi_j\right\|$都能够足够小, 那么, 函数系$\{\varphi_j(t)\}$就被称为是一个**完备正交系**. 正如在已往的推断中已经看到的, $\{\varphi_j(t)\}$的完备性的必要条件可以归结为, 对任意函数f都满足关系式

$$\sum_{j=1}^{\infty}|c_j|^2=\sum_{j=1}^{\infty}|(f|\varphi_j)^2=\|f\|^2 \tag{4}$$

(帕塞瓦尔不等式).

有关相对于给定函数类$f(t)$的完备标准正交系$\{\varphi_j(t)\}$的问题, 首先由著名的俄罗斯数学家B.A.斯捷克洛夫(B.A.Стеклов, 1864—1926)提出的.

完备性条件(4)可以表达成积分形式

$$\lim_{m\to\infty}\int_a^b\left(f(t)-\sum_{j=1}^{m}c_j\varphi_j(t)\right)^2dt=0,\quad \forall f\in C_2(a,b). \tag{5}$$

在满足条件(5)时, 还可以说是函数序列$\sum\limits_{j=1}^{m}c_j\varphi_j(t)$**按平均收敛**到$f(t)$. 一般说来, 按平均收敛不一定能推出按函数$\varphi_j(t)$分解成级数, 即$f(t)=\sum\limits_{j=1}^{\infty}c_j\varphi_j(t)$.只有在级数$\sum\limits_{j=1}^{\infty}c_j\varphi_j(t)$一致收敛的情况下, 据完备性条件(5)在积分符号下取极限函数, 同时把$f(t)$分解成级数的可能性才会变成现实.

按平均收敛的概念, 以及和它在一起的函数的完备系的概念, 对于不一定是正交的、标准的系仍然是有意义的.

3. 线性方程组与最小二乘法　与前面的说明有关, 而且同样可以得到对于最小二乘法的补充信息. 我们返回到由点到子空间的距离的计算问题.

照旧, 设V是个任意维数的带有纯量乘积$(*|*)$的向量空间, f是个固定的向量, 而$U=\langle \mathbf{e}_1,\cdots,\mathbf{e}_m\rangle$是$V$的一个子空间. 我们已经看出来, 由“点” f到空间U的距

离可以用向量$f-f_0$的模来度量, 其中f_0是U中以$x_1,\cdots,x_m$为坐标的向量, 而这些坐标可由线性方程组

$$x_1(\mathbf{e}_1|\mathbf{e}_j)+x_2(\mathbf{e}_2|\mathbf{e}_j)+\cdots+x_m(\mathbf{e}_m|\mathbf{e}_j)=(f|\mathbf{e}_j),\quad 1\leqslant j\leqslant m \tag{6}$$

来确定(见(3), 在该处用$\mathbf{e}_j$代替φ_j). 如果$\mathbf{e}_1,\cdots,\mathbf{e}_m$是标准正交组, 那么, $x_j=(f|\mathbf{e}_j)$, $1\leqslant j\leqslant m$, 就是待定解.但是, 解的存在性, 也就是有唯一确定的由f到U的垂直线, 却不自然地与某个基底有了关系. 我们知道, 任何基底总都可以化成标准正交基底, 所以(6)的解总是存在的, 而且是唯一的. 这意味着, 行列式

$$\det\|(\mathbf{e}_i|\mathbf{e}_j)\|_{1\leqslant i,j\leqslant m},$$

(也被称为向量组$\mathbf{e}_1,\cdots,\mathbf{e}_m$的**格拉姆行列式**)不等于零. 显然, 如果向量组中的某一个向量$\mathbf{e}_k$是其余向量的线性组合, 那么, 格拉姆行列式等于零. 事实上, 我们就证明了

定理1 向量组$\{\mathbf{e}_1,\cdots,\mathbf{e}_m\}$线性无关, 当且仅当它的格拉姆行列式不等于零.

这个断言不是新的(见§1 定理4). 与最小二乘法求近似问题一起, 似乎, 允许我们用全新的观点彻底地完成线性方程组的求解问题. 设给定一个$m>n$的线性方程组

$$\begin{aligned}
&a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\
&a_{21}x_1+a_{22}x_2+\cdots+a_{n2}x_n=b_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m.
\end{aligned} \tag{7}$$

不失一般性, 可以认为$\operatorname{rank}(a_{ij})=n$. 这种不定方程组是实践中可能发生的. 例如, 一般大面积计算时, 方程组(7)可能是不相容的, 也就是说, 它没有解. 但是, 可以尝试去找一组未知量的值x_1^0, x_2^0, $\cdots$, x_n^0使得平均二次误差

$$\sum_{k=1}^{m}\left(a_{k1}x_1^0+a_{k2}x_2^0+\cdots+a_{kn}x_n^0-b_k\right)^2$$

取极小值, 我们把列

$$\mathbf{e}_1=[a_{11},\cdots,a_{m1}],\cdots,\mathbf{e}_n=[a_{1n},\cdots,a_{mn}],\mathbf{f}=[b_1,\cdots,b_m]$$

解释成带有标准纯量乘积

$$([x_1,\cdots,x_m]\,|\,[y_1,\cdots,y_m])=\sum_{i=1}^{m}x_iy_i$$

的m维欧几里得空间的向量. 此时,

$$\sum_{k=1}^{m}\left(a_{k1}x_1^0+\cdots+a_{kn}x_n^0-b_k\right)^2=\left\|\sum_{i=1}^{n}x_i^0\mathbf{e}_i-\mathbf{f}\right\|^2$$

就是由$\sum_{i=1}^{n} x_i^0 \mathbf{e}_i$到$\mathbf{f}$的距离的平方. 如果$U = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle$是向量$\mathbf{e}_1, \cdots, \mathbf{e}_n$按线性无关条件的线性包络, 那么, 关于最小二乘法条件的问题就变成了已知的关于垂直的问题了, 也就是, 求向量$\mathbf{f}$到U上的投影问题. 我们关心的原来的线性方程组的“近似解”的分量x_1^0, $\cdots$, x_n^0刚好是在相容的确定的“正规的”方程组

$$
\begin{aligned}
&(\mathbf{e}_1|\mathbf{e}_1)\,x_1^0 + (\mathbf{e}_2|\mathbf{e}_1)\,x_2^0 + \cdots + (\mathbf{e}_n|\mathbf{e}_1)\,x_n^0 = (\mathbf{f}|\mathbf{e}_1),\\
&(\mathbf{e}_1|\mathbf{e}_2)\,x_1^0 + (\mathbf{e}_2|\mathbf{e}_2)\,x_2^0 + \cdots + (\mathbf{e}_n|\mathbf{e}_2)\,x_n^0 = (\mathbf{f}|\mathbf{e}_2),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&(\mathbf{e}_1|\mathbf{e}_n)\,x_1^0 + (\mathbf{e}_2|\mathbf{e}_n)\,x_2^0 + \cdots + (\mathbf{e}_n|\mathbf{e}_n)\,x_n^0 = (\mathbf{f}|\mathbf{e}_n)
\end{aligned}
$$

之中, 这个方程组的格拉姆行列式$\det \|(\mathbf{e}_i|\mathbf{e}_j)\| \neq 0$.

4. 三角多项式 由很容易验证的关系式,

$$
\begin{aligned}
&\int_{-\pi}^{\pi} \cos kt \cdot \cos lt\, dt = 0,\ k \neq l, \quad \int_{-\pi}^{\pi} \cos^2 kt\, dt = \pi,\\
&\int_{-\pi}^{\pi} \cos kt \cdot \sin lt\, dt = 0,\ k \neq l, \quad \int_{-\pi}^{\pi} 1 \cdot dt = 2\pi,\\
&\int_{-\pi}^{\pi} \sin kt \cdot \sin lt\, dt = 0,\ k \neq l, \quad \int_{-\pi}^{\pi} \sin^2 kt\, dt = \pi
\end{aligned}
$$

可以推出, 函数

$$
\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos t,\ \frac{1}{\sqrt{\pi}}\sin t,\ \cdots,\ \frac{1}{\sqrt{\pi}}\cos nt,\ \frac{1}{\sqrt{\pi}}\sin nt
$$

构成了$2n+1$维空间V_{2n+1}的一个标准正交基底, 而V_{2n+1}就是所谓的n阶三角函数多项式

$$
s_n(t) = \frac{a_0}{2} + a_1 \cos t + b_1 \sin t + \cdots + a_n \cos nt + b_n \sin nt \tag{8}
$$

的向量空间.

对于函数逼近的一般推理表明, 对任意函数$f \in C_2(-\pi, \pi)$, 带(似乎是可变的)傅里叶系数

$$
\begin{aligned}
&a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)dt,\\
&a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\, dt, \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt\, dt, \quad 1 \leqslant k \leqslant n,
\end{aligned}
$$

的三角多项式$S_n(t)$给出了按平均值(n阶)的最佳逼近.

已经发展极为成熟的傅里叶级数理论要对不同的函数类解决它们是不是能给出最佳平均值逼近的傅里叶多项式的问题, 而且还要求一致逼近, 也就是说, 无穷级数$\lim\limits_{n\to\infty} S_n(t)$是否一致收敛, 而且收敛到$f(t)$. 特别地, 对于在区间$[-\pi, \pi]$上的任意一

个具有自然条件$f(-\pi) = f(\pi)$的连续函数$f(t)$, 回答是肯定的. 这个结论可以由即将在第6目中描述的魏尔斯特拉斯定理推导出来, 我们不打算加以证明.

我们有一个重要的技术性的说明. 借助已经遇到过的欧拉公式$\cos kt + i\sin kt = e^{ikt}$, 对于复函数$f(t)$的三角函数多项式(8)就可以表达成更方便的形式

$$s_n(t) = \sum_{k=-n}^{n} \alpha_k e^{ikt} \tag{9}$$

(请注意从$-n$到n的不寻常的对称), 其中

$$\alpha_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt = \frac{1}{2\pi}(f|e^{ikt}),$$

$$2\alpha_0 = a_0; \quad 2\alpha_k = a_k - ib_k,\ \ k > 0; \quad 2\alpha_k = a_{-k} + ib_{-k},\ \ k < 0.$$

典型函数

$$\left\{\frac{1}{\sqrt{2\pi}}e^{int}|n \in \mathbb{Z}\right\}$$

给出了一个在区间$[-\pi, \pi]$上的复的标准正交函数系的例子, 这可以由正交性关系式

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-l)t}dt = \delta_{kl}.$$

直接得到.

5. 关于自共轭算子的说明 三角函数(典型函数)的标准正交系完备性给出了一个需要从某些不寻常方面来注视它们的道理. 事情在于, 很多完备的标准正交函数系都是作用在$C_2(a,b)$上或者某个子集合$\Omega \subset C_2(a,b)$上的某个自共轭算子的一整套的特征函数(特征向量).

一般来说, 当$\dim V < \infty$时, 自共轭算子$\mathcal{A}: V \to V$的可以对角化的性质并不能照搬到无穷维空间中来, 这只要举个以t相乘的算子$\mathcal{F}_t$的例子就可以说明白, 算子$\mathcal{F}_t$是对称的

$$(\mathcal{F}_t f(t)|g(t)) = \int_a^b tf(t)g(t)dt = \int_a^b f(t)tg(t)dt = (f(t)|\mathcal{F}_t g(t)).$$

但是, $\mathcal{F}_t e(t) = \lambda e(t) \Rightarrow e(t) \equiv 0$, 所以算子$\mathcal{F}_t$本身没有特征向量. 这可能会使我们想起, 在无穷维空间(即使是希尔伯特空间)上关于线性算子的其他的困难工作, 而这只是我们引到的一个方面.

更加重要的是要注意, 许多数学上和物理上具有重大意义的无穷维向量空间上的满足一系列条件的算子都是自共轭的. 而且, 对于它们而言, 一个自然地类似于有限维空间上的谱定理是正确的. 也就是说, 如果$\mathcal{A}: V \to V$是个自共轭算子, 那么, 它的特征向量(函数)的标准正交系$\mathcal{S}_\mathcal{A}$在第2目的意义下常常是完备的. 我们举一个简单而又通俗的例子来说明这个值得注意的事实.

进一步, 设$C_2''(a,b)$是二次连续可微函数的空间, 它带有通常的纯量乘积

$$(f|g) = \int_{-\pi}^{\pi} f(t)\overline{g(t)}dt.$$

考察实函数的集合

$$\Omega = \{f \in C_2''(-\pi,\pi) | f(-\pi) = f(\pi),\ \ f'(-\pi) = f'(\pi)\}$$

和以Ω为定义域的线性算子

$$\mathcal{A} = \frac{d^2}{dt^2} : C_2(-\pi,\pi) \to C_2(-\pi,\pi),$$

用分部积分法, 得出

$$\begin{aligned}(\mathcal{A}f(t)|g(t)) &= \int_{-\pi}^{\pi} f''(t)g(t)dt \\ &= f'(t)g(t)|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(t)g'(t)dt = -\int_{-\pi}^{\pi} f'(t)g'(t)dt \\ &= f(t)g'(t)|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(t)g'(t)dt = \int_{-\pi}^{\pi} f(t)g''(t)dt = (f(t)|\mathcal{A}g(t)).\end{aligned}$$

这表明, 在所做的假设之下, 算子$\mathcal{A}$是自共轭的(对称的).关于它的特征向量和特征值能说明什么呢? 设

$$\frac{d^2f(t)}{dt^2} = \lambda f(t), \quad f(t) \in \Omega.$$

函数族

$$M_k \cos kt + N_k \sin kt \tag{$\circ$}$$

就是这个定义在Ω上带有已经阐述过的计算限制的方程的对应特征值$\lambda = -k^2 (k = 0, 1, 2, \cdots)$的解. 如果还存在某个特征值, 那么, 找到了一个与所有三角级数函数正交的函数(自共轭算子的对应于不同特征值的特征向量的正交性), 而由三角函数系的完备性可以知道这是不可能的. 当$\lambda = -k^2$时, ($\circ$)式已经穷尽了所有的解.

这样一来, 下面的定理成立.

定理2 定义在区间$[-\pi,\pi]$上且满足条件$f(-\pi) = f(\pi),\ \ f'(-\pi) = f'(\pi)$的二次连续可微函数类的微分方程式

$$\frac{d^2}{dt^2}f(t) = \lambda f(t)$$

仅仅在$\lambda = -n^2\ (n = 0, 1, \cdots)$时有解. 每个$n$对应一个二维空间$\langle \cos nt, \sin nt \rangle$, 所有的解, $1, \cos t, \sin t, \cdots$组成$C_2(-\pi,\pi)$中的完备标准正交的函数系.

术语"微分方程式"在这里暂且能理解为"特征值和特征函数上的方程式"

$$\mathcal{D}f(t) = \lambda f(t),$$

其中,

$$\mathcal{D} = a_m(t)\frac{d^m}{dt^m} + a_{m-1}(t)\frac{d^{m-1}}{dt^{m-1}} + \cdots + a_1(t)\frac{d}{dt} \tag{*}$$

是作用在所有的在区间$[a,b]$充分光滑的函数组成的线性空间上的微分算子; 而且我们还假定, 如果$f(t)$是这个类中的函数, 那么$f^{(k)}(a) = f^k(b)$, $k = 0,1,\cdots,m-1$. 利用分部积分法若干次, 就得到

$$\mathcal{D}^* = \sum_{i=1}^{m}(-1)^i\frac{d^i}{dt^i}\circ a_i(t), \tag{**}$$

其中, 对于算子的记法$\frac{d^i}{dt^i}\circ a_i(t)$意味着, 把它作用到函数$f(t)$上就是开始先用$a_i(t)$ 去乘, 然后再对t求i次导数. 公式(**)定义了形式上的共轭微分算子的运算$\mathcal{D}\mapsto\mathcal{D}^*$, 称算子$\mathcal{D}$是(**形式上的**)**自共轭的**, 如果$\mathcal{D}^*=\mathcal{D}$. "形式上的"一词在这里可以理解为在定义中没有明显地指出来的空间上, $\mathcal{D}$实际上是个线性算子.

6. 勒让德多项式(球面多项式) 回顾一下, 在第4目中的魏斯特拉斯定理称:*在区间$a\leqslant t\leqslant b$上的任意连续函数$f(t)$都可以被这个区间上的t多项式一致逼近.* 换句话说, 对任意的正的$\varepsilon > 0$, 对于$f(t)$都能找到一个次数n充分大的多项式$a_0 + a_1t + \cdots + a_nt^n$使得

$$|a_0 + a_1t + \cdots + a_nt^n - f(t)| < \varepsilon, \quad a\leqslant t\leqslant b.$$

这个在分析学教程中早已证明了的定理, 既表明了单项式的无穷系$\{t^i\}_0^\infty$的完备性, 又表明了按$f(t)$建立的相应的标准正交系的傅里叶级数均值(或按模)收敛到$f(t)$. 为了得到这个函数系, 我们对单项式t^k施以格拉姆–施密特正交化过程. 这个过程中产生的一系列正交的、标准的多项式是唯一的, 如果把区间固定, 比方说, $-1\leqslant t\leqslant 1$, 并且约定, 每个多项式的最高项的系数都是正的.

但是, 正交多项式族$\{\varphi_n(t)\}$常常并没有被规范成具有积分条件$\|\varphi_n(t)\| = 1$. 不管怎样, 总可以有下列类型之1的局部性约定

1) $\varphi_n(t)$是n次标准多项式, 即, $\varphi_n^{(n)}(t) = n!$;

2) $\varphi_n(1) = 1$.

在任何情况下, 都会得到平行的向量(函数)系, 因为, 正交性条件同样可以记成

$$\int_{-1}^{1} t^k\varphi_n(t)dt = 0, \quad k = 0,1,\cdots,n-1.$$

然后就可以再设n=1,2,3,$\cdots$并选择型1)的正则性, 我们得到多项式的一个正交系

$$u_0(t) = 1,\quad u_1(t) = t,\quad u_2(t) = t^2 - \frac{1}{3},\quad u_3(t) = t^3 - \frac{3}{5}t, \cdots \tag{10}$$

这是二百多年前(1785)法国数学家勒让德联系位势理论问题首先研究的. 对于它们的一般公式, 晚些时候再求证, 现在, 称具有型2)的正则性: $P_n(1) = 1$的正交的多项式族

$$P_0(t) = 1,\quad P_n(t) = \frac{1}{2^n n!}\frac{d^n(t^2-1)^n}{dt^n},\quad n = 1,2,\cdots \tag{11}$$

为勒让德多项式.

我们来验证, 事实上, 多项式族(11)具有所需要的性质.按照牛顿二项式定理, 我们有

$$(t^2-1)^n=\sum_{k=0}^{n}(-1)^k\binom{n}{k}t^{2(n-k)}=t^{2n}-nt^{2n-2}+\cdots$$

所以,

$$\begin{aligned}P_n(t)&=\frac{1}{2^n n!}[2n(2n-1)\cdots(n+1)t^n+\text{次数}\leqslant n-2\text{的项}\\&=\frac{(2n)!}{2^n(n!)^2}t^n+\text{次数较低的项}.\end{aligned}\tag{12}$$

这说明, $\deg\ P_n(t)=n$, 而且, 我们还得到了多项式$P_n(t)$的最高项的表达式.

其次, 把求乘积的n次微商的莱布尼茨公式用到多项式$(t^2-1)^n=(t-1)^n(t+1)^n$上, 我们就得到

$$\frac{d^n}{dt^n}[(t-1)^n(t+1)^n]=\sum_{k=0}^{n}\binom{n}{k}\frac{d^k}{dt^k}(t-1)^n\cdot\frac{d^{n-k}}{dt^{n-k}}(t+1)^n.$$

因为, 当$k<n$ 时, 多项式$\frac{d^k}{dt^k}(t-1)^n$一定能被$t-1$整除, 从而, 当n=1时, 它就变成零, 所以

$$P_n(1)=\frac{1}{2^n n!}\binom{n}{n}\left[\frac{d^n}{dt^n}(t-1)^n\right](t+1)^n|_{t=1}=\frac{1}{2^n n!}\cdot 1\cdot n!\cdot 2^n=1.$$

顺便说一句, 我们应当注意, 当$m<n$时, 求乘积$(t-1)^n(t+1)^n$的m 次微商的莱布尼茨公式给我们做出来一个既能被$t-1$整除, 又能用t+1整除的多项式, 即

$$\frac{d^m}{dt^m}(t^2-1)^n=(t^2-1)\cdot v_m(t),\quad m<n.$$

可见, 多项式

$$\frac{d^m}{dt^m}(t^2-1)^n,\quad m<n,$$

在±1处均为零. 现在, 利用分部积分公式, 就可以验证$P_n(t)$与$1,\ t,\cdots,t^{n-1}$的正交性. 我们有

$$\begin{aligned}2^n n!(t^k|P_n(t))&=\int_{-1}^{1}t^k\frac{d^n}{dt^n}(t^2-1)^n dt\\&=t^k\frac{d^{n-1}}{dt^{n-1}}(t^2-1)^n|_{-1}^{+1}-k\int_{-1}^{1}t^{k-1}\frac{d^{n-1}}{dt^{n-1}}(t^2-1)^n dt\end{aligned}$$

$$= -kt^{k-1}\frac{d^{n-2}}{dt^{n-2}}(t^2-1)^n|_{-1}^{+1} + k(k-1)\int_{-1}^{1} t^{k-2}\frac{d^{n-2}}{dt^{n-2}}(t^2-1)^n dt.$$

逐次地降低t的幂指数, 就得到最终的等式

$$2^n n!(t^k|P_n(t)) = (-1)^k k!\frac{d^{n-k-1}}{dt^{n-k-1}}(t^2-1)^n\,|_{-1}^{+1} = 0.$$

用这种间接的方法, 我们不仅证明了勒让德多项式的两两正交性

$$(P_k(t)|P_l(t)) = 0, \quad k \neq l,$$

而且得到了序列(10)的多项式表达式

$$u_n(t) = \frac{2^n(n!)^2}{(2n)!}P_n(t), \quad n = 1, 2, \cdots.$$

实际上, 由一般的表达式, 我们知道$u_n(t)$和$P_n(t)$仅仅差一个常数因子, 而且, 比较它们的最高项系数(见(12)式)就可以做出低项的关系式.

可以验证,

$$\|P_n(t)\|^2 = \int_{-1}^{1} P_n^2(t)dt = \frac{2}{2n+1}. \tag{13}$$

我们再设$w_n = (t^2-1)^n$.借助莱布尼茨公式对恒等式

$$(t^2-1)\frac{d}{dt}w_n = 2ntw_n,$$

两端的乘积求$n+1$次导数, 就得到

$$(t^2-1)\frac{d^{n+2}}{dt^{n+2}}w_n + 2(n+1)t\frac{d^{n+1}}{dt^{n+1}}w_n + (n+1)n\frac{d^n}{dt^n}w_n$$
$$= 2nt\frac{d^{n+1}}{dt^{n+1}}w_n + 2n(n+1)\frac{d^n}{dt^n}w_n.$$

把这个等式的所有项均乘以$1/(2^n n!)$且采用

$$\frac{1}{2^n n!}\frac{d^n}{dt^n}w_n = P_n(t),$$

我们就得到一个微分关系式

$$(t^2-1)\frac{d^2}{dt^2}P_n(t) + 2t\frac{d}{dt}P_n(t) - n(n+1)P_n(t) = 0. \tag{14}$$

在空间$C_2(-1,1)$上研究线性微分算子

$$\mathcal{S} = (t^2-1)\frac{d^2}{dt^2} + 2t\frac{d}{dt} = \frac{d}{dt}\left[(t^2-1)\frac{d}{dt}\right],$$

它的定义域是$C_2''(-1,1)$, 其中, 如同从前一样(见第5目), $C''(a,b)$是带有通常纯量乘积的所有的二次连续可微函数的空间.算子$\mathcal{S}$是自共轭的, 这可以由第5目末尾的一般公式(*)和(**)直接推出来, 可以改写成

$$\mathcal{S}P_n(t) = n(n+1)P_n(t), \tag{15}$$

形式的原来的等式(14)表明, 多项式$P_n(t)$是自共轭算子$\mathcal{S}$的属于特征值$\lambda = n(n+1)$的一个特征函数. 换句话来说, 方程式$\mathcal{S}x(t) = n(n+1)x(t)$有非零解$x(t) = P_n(t)$. 如果特征子空间$V^\lambda$的维数能够大于1, 那么, V^λ中就会有向量$y(t) \neq 0$, 它正交于$P_n(t)$. 因为自共轭算子的特征子空间相互正交:

$$(V^\lambda|V^\mu) = 0, \quad \lambda \neq \mu,$$

因此, 向量$y(t)$就正交于所有的$P_j(t)$, $j = 0,1,\cdots$. 但这与由魏尔斯特拉斯推出的系(11)的完备性相矛盾. 出于同样的原因, 算子$\mathcal{S}$必然没有不同于$n(n+1)$, $n = 0,1,\cdots$ 的特征值.

这样, 就证明了(按魏尔斯特拉斯定理的模式)下面的断言

定理3　*微分方程*

$$\frac{d}{dt}\left[(t^2-1)\frac{d}{dt}x(t)\right] = \lambda x(t)$$

在定义于区间$-1 \leqslant t \leqslant 1$上的所有二次连续可微函数族中, 只有当$\lambda = n(n+1)$, n=0,1,2,$\cdots$时, 有解. 每个n对应唯一一个解$x(t) = P_n(t)$(精确到乘一个常数). 上述的所有解在$C_2(-1,1)$中构成一个完备的正交函数系.

我们研究过(极其短暂地)的微分算子$\dfrac{d^2}{dt^2}$和$\dfrac{d}{dt}(t^2-1)\dfrac{d}{dt}$都属于一个更广泛的算子类, 被称为**斯图姆–刘维尔算子类**, 这类算子在数学物理中起卓越作用.

说明　系(10)中多项式

$$u_n(t) = \alpha_n P_n(t), \quad \alpha_n = \frac{2^n(n!)^2}{(2n)!}$$

具有如下有趣的极小性. 在所有的标准n次实多项式中, 至少$u_n(t)$在区间$-1 \leqslant t \leqslant 1$上按均值不为零.

事实上, 问题是积分

$$I(f) = \int_{-1}^{1} f(t)^2 dt \quad \text{其中} \quad f(t) = t^n + \cdots \in \mathbb{R}[t]$$

的极小性. 把多项式$P_k(t)$的两两正交性和对于$\| P_n(t) \|$ 的公式(13)应用到分解式

$$f(t) = \alpha_n P_n(t) + \gamma_{n-1}P_{n-1}(t) + \cdots + \gamma_1 P_1(t) + \gamma_0, \quad \gamma_i \in \mathbb{R}$$

上, 我们就得到表达式

$$I(f) = \frac{2\alpha_n^2}{2n+1} + 2\sum_{i=0}^{n-1}\frac{\gamma_i^2}{2i+1},$$

显然, 在$\gamma_i = 0, 0 \leqslant i \leqslant n-1$时, 上面的积分取得极小值.

7. 加权正交　以下述方式得到的为数众多的函数族是勒让德多项式的直接推广. 设在区间$a \leqslant t \leqslant b$上给定了一个非负函数$p(t)$, 我们称它为**权函数**. 考察向量空间

$$V(\sqrt{p(t)}) = \left\langle \sqrt{p(t)}t^k | k = 0, 1, 2, \cdots \right\rangle_{\mathbb{R}}$$

或者它的有限维子空间

$$V_n(\sqrt{p(t)}) = \left\langle \sqrt{p(t)}t^k | 1 \leqslant k \leqslant n-1 \right\rangle_{\mathbb{R}}.$$

我们眼前就呈现了关于在$V(\sqrt{p(t)})$(或者$V_n(\sqrt{p(t)})$)中选取标准正交基底的问题. 把通常的格拉姆–施密特正交化过程用到这个函数类, 得到满足条件

$$(\sqrt{p(t)}Q_m(t)|\sqrt{p(t)}Q_n(t)) = \int_a^b p(t)Q_m(t)Q_n(t)dt = \delta_{mn}$$

的函数类

$$\left\{\sqrt{p(t)}Q_n(t)\right\}, \quad Q_n(t) \in \mathbb{R}[t], \quad \deg Q_n = n, \quad n = 0, 1, \cdots.$$

可以说$\{Q_n(t)\}$是**相应于权$p(t)$的标准正交多项式**.在这个意义下, 勒让德多项式就相应于权1的标准正交多项式.

我们本来在一开始就引入新的纯量乘积

$$(f|g)_{p(t)} = \int_a^b p(t)f(t)g(t)dt,$$

那么, 关于前面意义之下的多项式正交化问题, 一如既往, 可认定与固定的纯量乘积有关系.

8. (第一类)切比雪夫多项式　具有多方面兴趣的俄罗斯数学家和力学家П.Л. 切比雪夫(1821—1894)为函数逼近理论奠定了基础. 他提出了正交多项式理论的基本思想, 由权$p(t) = 1/\sqrt{1-t^2}$对应的极有特色的标准正交多项式类$T_n(t), n \geqslant 0, (a, b) = (-1, 1)$就是用他的名字命名的. 下面就是它们的显式

$$T_n(t) = \frac{(-2)^n n!}{(2n)!}\sqrt{1-t^2}\frac{d^n}{dt^n}(1-t^2)^{n-1/2} = \cos(n \arccos t). \tag{16}$$

特别地,

$$T_0(t) = 1,\ T_1(t) = t,\ T_2(t) = 2t^2 - 1,$$
$$T_3(t) = 4t^3 - 3t,\ T_4(t) = 8t^4 - 8t^2 + 1.$$

规范化:

$$\int_{-1}^1 \frac{T_m(t)T_n(t)dt}{\sqrt{1-t^2}} = \begin{cases} 0, & 当 \quad m \neq n, \\ \pi/2, & 当 \quad m = n \neq 0, \\ \pi, & 当 \quad m = n = 0. \end{cases}$$

就第6目末尾的说明, 我们看出, 切比雪夫多项式在如下的意义上偏离零最少: 在闭区间$-1 \leqslant t \leqslant 1$上绝对值$\left|\frac{1}{2^{n-1}}T_n(t)\right|$的极大值在所有标准的$n$次实多项式的类中取值最小.

习 题

验证, 表达式(16)是对的, 并且证明, 切比雪夫多项式$T_n(t)$是微分算子

$$(t^2-1)\frac{d^2}{dt^2}+t\frac{d}{dt}$$

的一个属于特征值n^2的特征向量.

9. 埃尔米特多项式 我们简短地考察一下埃尔米特多项式$H_n(t)$ (更规范地应该说是拉普拉斯–切比雪夫–埃尔米特多项式), 它是先分别选取$a=-\infty$, $b=\infty$, $p(t)=e^{-t^2}$, 然后在单项式序列1, t, $t^2, \cdots$ 正交化基底上得到的. 显式是

$$H_n(t)=(-1)^n e^{t^2}\frac{d^n}{dt^n}e^{-t^2};$$
$$H_0(t)=1,\ H_1(t)=2t,\ H_2(t)=4t^2-2,\ H_3(t)=8t^3-12t,\cdots;$$

规范化:

$$\int_{-\infty}^{\infty}e^{-t^2}H_m(t)H_n(t)dt=\begin{cases}0, & 当\quad m\neq n,\\ 2^n n!\sqrt{\pi}, & 当\quad m=n.\end{cases}$$

这对读者来说是个不难的习题, 只要掌握了第5目的内容就行.这只需要利用反常积分

$$\int_{-\infty}^{\infty}e^{-t^2}dt=\sqrt{\pi}$$

的值以及在区间$(-\infty,\infty)$终点函数e^{-t^2}的所有乘积均趋于零的情况.

其次, 对n直接用归纳法, 即可建立如下的命题.

埃尔米特多项式$H_n(t)$是微分算子

$$\mathcal{K}=\frac{d^2}{dt^2}-2t\frac{d}{dt}$$

的属于特征值$-2n$的一个特征向量.

在数学物理中, 所谓的**埃尔米特函数**

$$\psi_n(t)=e^{-t^2/2}H_n(t)=(-1)^n e^{t^2/2}\frac{d^n}{dt^n}e^{-t^2}$$

同样是很有用的.

我们来证明, 函数$\psi_n(t)$是算子

$$\mathcal{H}=\frac{d^2}{dt^2}-t^2$$

的属于特征值$-(2n+1)$的一个特征向量.

为了这个目的, 考察一个辅助算子

$$\mathcal{M}=\frac{d}{dt}-t.$$

容易确信,

$$[\mathcal{H},\mathcal{M}]=\mathcal{H}\mathcal{M}-\mathcal{M}\mathcal{H}=-2\left(\frac{d}{dt}-t\right)=-2\mathcal{M}.$$

由此可以得出, 如果f是算子$\mathcal{H}$的属于特征值λ的特征向量, 那么, $\mathcal{M}f$就是算子$\mathcal{H}$的属于特征值$\lambda-2$的特征函数:

$$\mathcal{H}\mathcal{M}f=[\mathcal{H},\mathcal{M}]f+\mathcal{M}\mathcal{H}f=-2\mathcal{M}f+\lambda\mathcal{M}f=(\lambda-2)\mathcal{M}f.$$

对n用归纳法, 可以推出关系式

$$\mathcal{H}\mathcal{M}^nf=(\lambda-2n)\mathcal{M}^nf.$$

因为$\mathcal{H}e^{-t^2/2}=-e^{-t^2/2}$, 那么, 用$e^{-t^2/2}$替换$f$, 用$-1$替换$\lambda$, 我们就可以推出这样一个事实, 即对所有的$n\geqslant 0, \mathcal{M}^ne^{-t^2/2}$乃是算子$\mathcal{H}$的属于特征值$-(2n+1)$的一个特征函数.

另一方面, 按n归纳, 可以直接计算出

$$e^{t^2/2}\frac{d^n}{dt^n}e^{-t^2}=\mathcal{M}^ne^{-t^2/2},$$

也就是

$$\psi_n(t)=(-1)^n\mathcal{M}^ne^{-t^2/2}.$$

这就是所有要证明的.

习　　题

1. 三角函数序列$\sum\limits_{k=1}^{\infty}\dfrac{\sin kt}{\sqrt{k}}$在$\mathbb{R}$上是收敛的. 利用贝塞尔不等式证明, 它不是任何一个函数$f\in C_2(-\pi,\pi)$的傅里叶级数.

2. 证明递推公式:

(1) 对勒让德多项式, $P_{n+1}(t)=\dfrac{2n+1}{n+1}tP_n(t)-\dfrac{n}{n+1}P_{n-1}(t)$;

(2) 对切比雪夫多项式, $T_{n+1}(t)=2tT_n(t)-T_{n-1}(t)$;

(3) 对埃尔米特多项式, $H_{n+1}(t)=2tH_n(t)-2nH_{n-1}(t)$.

3. 不必依据第5目中的公式(**)而直接证明第6目中的微分算子$\mathcal{S}$的自共轭性.

4. 证明, $f(x,t)=\dfrac{1}{\sqrt{1-2tx+x^2}}$是对勒让德多项式的生成函数(把$P_n(t)$当做是$f(x,t)$按幂$x^n$分解时的系数).

5. 证明, $\max\limits_{-1\leqslant t\leqslant 1}\left|\dfrac{1}{2^{n-1}}T_n(t)\right|=\dfrac{1}{2^{n-1}}$.

6. 证明, 第二类切比雪夫多项式

$$U_n(t) := \frac{1}{n+1}\frac{dT_{n+1}(t)}{dt} = \sum_{m=0}^{[n/2]} \binom{n+1}{2m+1} t^{n-2m}(t^2-1)^m$$

带上权函数$\sqrt{1-t^2}$在区间$[-1,1]$上是正交的:

$$\int_{-1}^{1} \sqrt{1-t^2}U_m(t)U_n(t)dt = \begin{cases} 0, & m \neq n, \\ \pi/2, & m = n. \end{cases}$$

7. 证明, 多项式$T_n(t)$, $U_n(t)$的所有的根都是实的, 且两两不同, 同时还都位于区间$[-1,1]$之内部. 试把这些根表达成显形式.

第4章　仿射空间与欧几里得点空间

对于生活在3维物理世界(用$\mathbb{R}^3_\Phi$ 代表)的我们来说, 离不开那些彼此之间放置成稀奇古怪形状的点, 直线, 平面, 它们不依赖于挑选出来的一个点——坐标原点打交道. 显然, 一般情况下, 最好是用位移方式来研究由过坐标原点的直线和平面生成的几何对象. 换句话说, 暂且离开度量, 我们希望能够把向量空间的自同构群加以扩展, 使得所有的向量(现在, 已经是"仿射空间"中的点)都等价, 向量空间是齐次的. 在这一章将引进所有必要的概念并证明仿射空间的一些最简单的性质.

§1　仿射空间

1. 仿射空间的定义　正如上面已经指出过的, 在任何一个向量空间中, 与零向量结合在一起的坐标原点都起着特殊的作用: 在空间的所有自同构之下, 零向量总是保持不变. 只能借助空间中的平移(平行的移动)把一般线性群推广之后, 所有的向量才能是平等的(等价的). 为了使这种思考获得准确的意义, 我们引进一些定义.

定义1　设$\mathbb{A}$ 是一个非空集合, 把它的元素叫做**点**, 并且用[1]$\dot{p}$, $\dot{q}$, $\dot{r}, \cdots$ 来代表.

其次, 设V 是某个域$\mathcal{K}$上的向量空间. 把集合$\mathbb{A}$ (更准确地, 对($\mathbb{A}$, V))称为是和V**相伴的**(连带的)**仿射空间**, 如果从给定的笛卡儿积$\mathbb{A}\times V$ 到$\mathbb{A}$ 的映射, $(\dot{p}, \mathbf{v}) \mapsto \dot{p}+\mathbf{v}$有如下的性质:

i) 对任意$\dot{p}\in\mathbb{A}$ 及任意向量$\mathbf{u}$, $\mathbf{v}\in V$ 都有$\dot{p}+\mathbf{0}=\dot{p}$, $(\dot{p}+\mathbf{u})+\mathbf{v}=\dot{p}+(\mathbf{u}+\mathbf{v})$ ($\mathbf{0}$是空间V 的零向量);

---

1) 经常地, 特别是在力学和微分方程理论中, 符号$\dot{p}$ 代表对可微函数$p=p(t)$求导数dp/dt. 在这里不会遇到这种情况.

ii) 对任意两点$\dot{p}$,$\dot{q}$ 都能找到唯一的一个向量$\mathbf{v} \in V$ 使得$\dot{p}+\mathbf{v}=\dot{q}$ (这个"由$\dot{p}$ 到$\dot{q}$ 的向量"通常用$\overrightarrow{pq}$或者$\dot{q}-\dot{p}$来代表).

就把向量空间V 的维数$n=\dim_{\mathfrak{K}} V$ 设定为与V 相伴的仿射空间$\mathbb{A}$ 的维数, 为了强调维数的作用, 有时写成$\mathbb{A}^n$. 在科学与技术中, 最令人感兴趣的是$\mathfrak{K}=\mathbb{R}$ 和$\mathfrak{K}=\mathbb{C}$ 的情形, 也就是实的或者相应地复的仿射空间.

按公理ii)本身的意思可以推出, 每个点$\dot{p} \in \mathbb{A}$ 都对应一个集合之间的双射$\mathbf{v} \to \dot{p}+\mathbf{v}\ : V \cong \mathbb{A}$. 另一方面, 我们有集合$\mathbb{A}$ 上的双射映射

$$t_{\mathbf{v}} : \dot{p} \to \dot{p}+\mathbf{v}=t_{\mathbf{v}}(\dot{p}), \quad \dot{p} \in \mathbb{A},$$

并称它是用向量$\mathbf{v}$ **平移**$\mathbb{A}$ (或者平行移动$\mathbb{A}$). 由公理i), ii)可以得到

$$t_{\mathbf{u}} \cdot t_{\mathbf{v}}=t_{\mathbf{u}+\mathbf{v}}, \qquad t_{\mathbf{v}} \cdot t_{-\mathbf{v}}=e$$

($e:=t_0$是恒等映射), 也就是说, $t_{-\mathbf{v}}$ 也是个平移, 而且是$t_{\mathbf{v}}$ 的逆映射. 可见, 所有的平移构成一个群, 它同构于空间V 的加法群. 如果令

$$\alpha t_{\mathbf{u}}+\beta t_{\mathbf{v}}:=t_{\alpha \mathbf{u}+\beta \mathbf{v}},$$

那么, 所有的平移的集合就构成了一个向量空间. 它是由空间$\mathbb{A}$ 唯一确定的而且同构于V , 用符号$\mathbb{A}^{\#}$ 代表它.

说明 要注意这样一个情况, 在意义完全不同的表达式$\mathbf{u}+\mathbf{v}$, $\dot{p}+\mathbf{v}$ 中用了同一个加号+, 但是, 这不致于引起误解. 其次, 如果有$\mathbb{A}$ 中的一些点$\dot{p}$, $\dot{q}$, $\dot{r}$, $\dot{s}$使得$\dot{p}+\mathbf{v}=\dot{q}$, $\dot{r}+\mathbf{v}=\dot{s}$, 那么, $\overrightarrow{pq}, \overrightarrow{rs}$作为等价类的不同的代表元素没有别的什么区别, 这个等价类就用向量$\mathbf{v}$ 来代表. 记法$\dot{p}+\overrightarrow{pq}=\dot{q}$ 使用起来很方便, 有利于记忆. 由定义可以直接得到一些用向量$\overrightarrow{pq}$作用的规则

$$\overrightarrow{pq}+\overrightarrow{qr}=\overrightarrow{pr}, \quad \overrightarrow{pq}=-\overrightarrow{qp}, \quad \overrightarrow{pp}=\mathbf{0}$$

($\dot{p}$, $\dot{q}$, $\dot{r}$ 是$\mathbb{A}$ 中的任意点). 由这些规则可以写出

$$(\dot{q}-\dot{p})+(\dot{r}-\dot{q})=\dot{r}-\dot{p}, \quad (\dot{q}-\dot{p})=-(\dot{p}-\dot{q}), \quad \dot{p}-\dot{p}=\mathbf{0}.$$

例1 如果V 是域$\mathfrak{K}$上的任意一个向量空间, 而$\mathbb{A}=\mathbf{v}_0+U$是对子空间$U \subset V$ 的一个陪集($\mathbf{v}_0$ 是V 中的一个固定的向量). 那么, $\mathbb{A}$ 就是$\mathfrak{K}$上的一个仿射空间, 而且平移空间$\mathbb{A}^{\#}=U$. 每一个向量$\mathbf{u}' \in U$ 都对应一个双射$\mathbf{v}_0+\mathbf{u} \mapsto \mathbf{v}_0+\mathbf{u}+\mathbf{u}'$, 这些个双射满足公理i), ii), 因为空间V 按加法构成一个群. 可以说, $\mathbb{A}$ 是空间V 的一个**仿射线性流形**(或者, **线性流形**), 而子空间是线性流形$\mathbb{A}$ 的**方向**.

特别地, 当$U=V$, 进而$\mathbb{A}$ 作为一个集合与V 重合的时候, 就记成$V_a:=\mathbb{A}$, 把点$\dot{p} \in V_a$ 直接理解为某个向量$\mathbf{u} \in V$. 于是, 对任意$\mathbf{v} \in V$, 有$\dot{p}+\mathbf{v}=\mathbf{u}+\mathbf{v} \in V_a$,

且映射$V_a \times V \to V_a$具有性质i), ii). 这个时候, $(V_a)^{\#} \cong V$. 实际上, 在同一个集合V上定义了两个不同的代数结构.

2. 同构 在同一个域$\mathcal{K}$上与同一个向量空间V相伴的两个仿射空间$\mathbb{A}$, $\mathbb{A}'$自然地被称为**同构的**, 如果存在某个双射映射f: $\mathbb{A} \to \mathbb{A}'$使得对任意$\mathbf{v} \in V$, $\dot{p} \in \mathbb{A}$都有$f(\dot{p}+\mathbf{v}) = f(\dot{p})+\mathbf{v}$ (为了简捷, 我们用同一种符号表示把平移$t_{\mathbf{v}}$作用到$\mathbb{A}$和$\mathbb{A}'$上得到的结果).

给出更一般的

定义2 设$\mathbb{A}$, $\mathbb{A}'$是同一个域$\mathcal{K}$上分别与向量空间V, V'相伴的仿射空间. 称映射$f: \mathbb{A} \to \mathbb{A}'$是个**仿射映射**(或者**仿射线性映射**), 如果, 所有的$\dot{p} \in \mathbb{A}$ $\mathbf{v} \in V$都满足关系式

$$f(\dot{p}+\mathbf{v}) = f(\dot{p}) + Df \cdot \mathbf{v}, \tag{1}$$

其中, $Df: V \to V'$是向量空间上的线性映射. 有时, 也称映射Df是映射f的**线性部分**(或者**微分**). 对于一个双射的仿射线性映射f, 它的线性部分Df同样是双射的. 在这种情形, 就说$\mathbb{A}$和$\mathbb{A}'$之间是**同构的**, 而当$\mathbb{A}=\mathbb{A}'$时, 就称空间$\mathbb{A}$是借助于非退化仿射变换f实现的自同构.

注意, 在前面所说的一般表示$\dot{p}+\mathbf{v}=\dot{q}$中方程式(1)可以写成

$$Df \cdot \overrightarrow{pq} = \overrightarrow{f(p)f(q)}. \tag{1'}$$

定理1 具有相同维数的仿射空间$(\mathbb{A}, V)$, $(\mathbb{A}', V')$必然同构.

证明 因为$\dim V = \dim \mathbb{A} = \dim \mathbb{A}' = \dim V'$, 故存在双射的线性映射$\mathcal{F}$: $V \to V'$ (第1章§2的定理5). 取固定点$\dot{o} \in \mathbb{A}$和$\dot{o}' \in \mathbb{A}'$, 我们只要令$f(\dot{o}) = \dot{o}'$, $Df = \mathcal{F}$, 就完全定义了$f: \mathbb{A} \to \mathbb{A}'$. 任意一点$\dot{p} \in \mathbb{A}$都可以写成$\dot{p} = \dot{o}+\mathbf{v}$. 按照我们的定义, 有

$$f(\dot{p}) = \dot{o}' + \mathcal{F}(\mathbf{v}). \tag{2}$$

当$\dot{p}$取遍$\mathbb{A}$的所有点, $\mathbf{v}$取遍V的所有向量(按照仿射空间的定义)时, 由于$\mathcal{F}$是个双射, $\dot{o}'+\mathcal{F}(\mathbf{v})$就取遍了$\mathbb{A}'$的所有点. 同样地, 我们使得$\mathbb{A}$中的不同点对应于$\mathbb{A}'$中的不同点. 可见, f是个双射. 剩下需要验证的是它的线性性质. 事实上, 应用(2), 我们得到

$$\begin{aligned} f(\dot{p}+\mathbf{u}) &= f((\dot{o}+\mathbf{v})+\mathbf{u}) = f(\dot{o}+(\mathbf{v}+\mathbf{u})) \\ &= \dot{o}' + \mathcal{F}(\mathbf{v}+\mathbf{u}) = \dot{o}' + (\mathcal{F}(\mathbf{v}) + \mathcal{F}(\mathbf{u})) \\ &= (\dot{o}' + \mathcal{F}(\mathbf{v})) + \mathcal{F}(\mathbf{u}) = f(\dot{p}) + Df \cdot \mathbf{u}. \end{aligned}$$

□

3. 坐标 我们引入一个很自然的

定义3 称点$\dot{o} \in \mathbb{A}$和V的一个基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$的集合$\{\dot{o}; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$是$n$维仿射空间$(\mathbb{A}, V)$的一个**坐标系**(或者**坐标架**). 向量$\overrightarrow{op}$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下:

$\overrightarrow{op} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ 的坐标就被认为是点$\dot{p}$ 在坐标系$\{\dot{o};\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ 之下的**坐标**.

由等式$\overrightarrow{pq} = \overrightarrow{oq} - \overrightarrow{op}$ 可以推出, 如果$x_1, \cdots, x_n$ 是点$\dot{p}$ 的坐标, 而$y_1, \cdots, y_n$ 是点$\dot{q}$ 的坐标, 那么$\overrightarrow{pq}$在$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的坐标就是$x_1 - y_1, \cdots, x_n - y_n$. 反过来, 如果$\dot{q} = \dot{p} + \mathbf{a}$, 那么, 点$\dot{q}$ 的坐标$y_1, \cdots, y_n$ 可以把向量的坐标$a_1, \cdots, a_n$ 与点$\dot{p}$ 的坐标$x_1, \cdots, x_n$ 相加而得到: $y_i = a_i + x_i,\ i = 1, \cdots, n$.

说明 坐标系也可以用这样的$n+1$ 个点$\{\dot{p}_0;\ \dot{p}_1, \cdots, \dot{p}_n\}$ 给出, 只要向量$\overrightarrow{p_0p_1}, \cdots, \overrightarrow{p_op_n}$ 构成空间V 的一个基底即可.

上述关于坐标的基本运算、表达方式, 可以归纳为

定理2 设$\{\dot{p}_0;\ \dot{p}_1, \cdots, \dot{p}_n\}$ 是空间V 的一个坐标系, $\mathbf{e}_i := \overrightarrow{p_0p_i}, i = 1, \cdots, n$. 如果在这个系统里, 点$\dot{p}$, $\dot{q}$ 的坐标分别是$x_1, \cdots, x_n$ 和$y_1, \cdots, y_n$, 那么, 向量$\overrightarrow{pq}$在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的坐标就是$y_1 - x_1, \cdots, y_n - x_n$. 对任意向量$\mathbf{a} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$, 点$\dot{p} + \mathbf{a}$ 的坐标是$x_1 + a_1, \cdots, x_n + a_n$.

假定我们现在想从坐标系$\{\dot{o};\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ 向坐标系$\{\dot{o}';\mathbf{e}'_1, \cdots, \mathbf{e}'_n\}$ 转化, 那么, 就需要给出$\dot{o}'$ 在原来坐标系的坐标$b_1, \cdots, b_n$ (也就是向量$\overrightarrow{oo'}$的坐标)以及从空间V 的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$到基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$的转换矩阵$A = (a_{ij})$ (图5). 设$x_1, \cdots, x_n$和$x'_1, \cdots,$

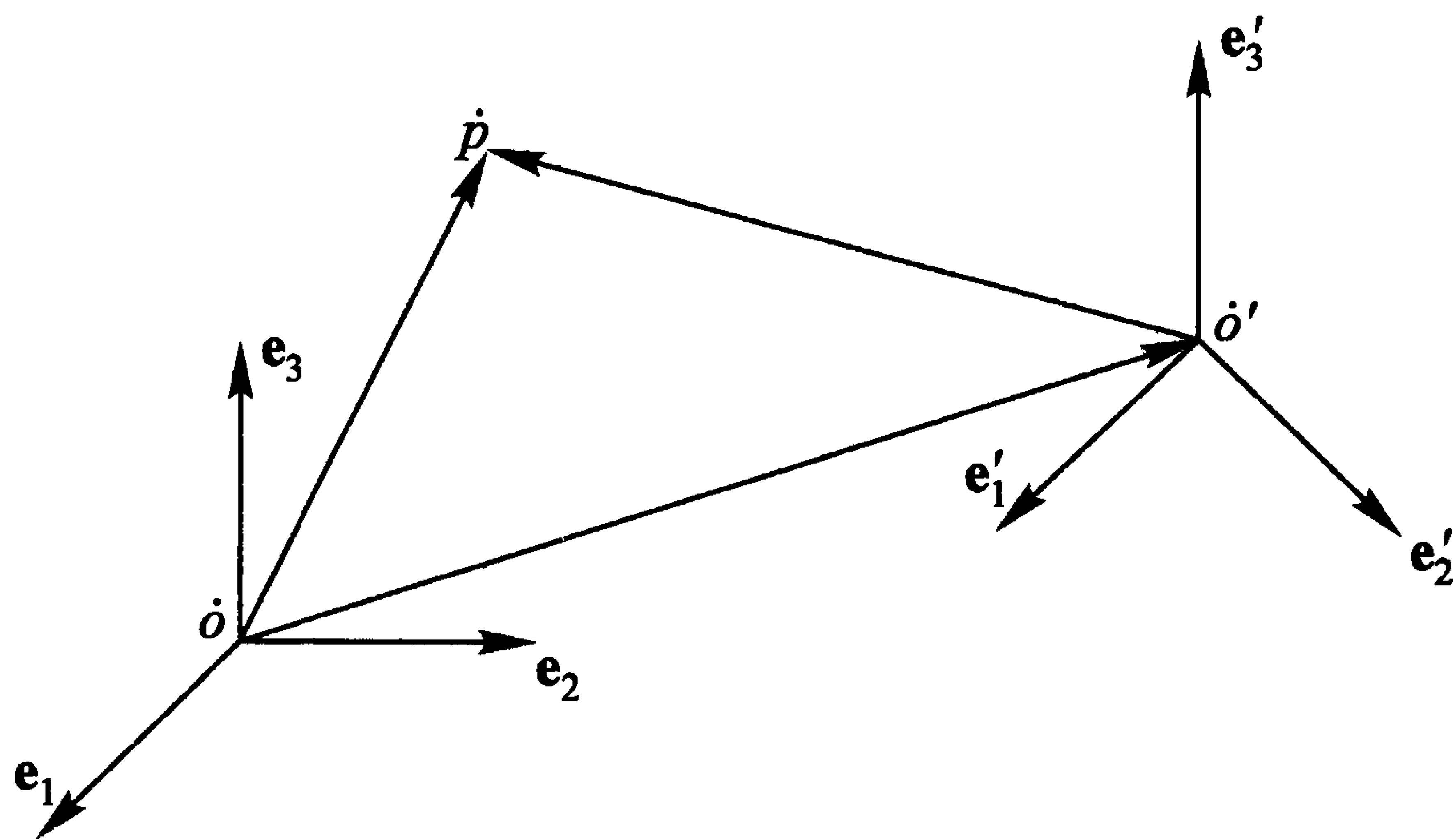


图5

x'_n 分别是点$\dot{p} \in \mathbb{A}$ 的老坐标与新坐标. 由等式

$$\overrightarrow{op} = \overrightarrow{oo'} - \overrightarrow{po'} = \sum_i b_i\mathbf{e}_i + \sum_j x'_j\mathbf{e}'_j = \sum_i b_i\mathbf{e}_i + \sum_j x'_j \sum_i a_{ij}\mathbf{e}_i$$

$$= \sum_i \left(\sum_j a_{ij}x'_j \right) \mathbf{e}_i + \sum_i b_i\mathbf{e}_i$$

可以推出

$$x_i = \sum_{j=i}^{n} a_{ij}x'_j + b_i, \quad i = 1, \cdots, n. \tag{3}$$

也就是

$$X = AX' + B,$$

其中

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad X'=\begin{pmatrix}x'_1\\ \vdots\\ x'_n\end{pmatrix},\quad B=\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}.$$

因为det $A\neq 0$, 所以

$$X'=A^{-1}X+B',\quad B'=\begin{pmatrix}b'_1\\ \vdots\\ b'_n\end{pmatrix}=A^{-1}B.$$

4. 仿射子空间 为了进一步的学习, 我们引进

定义4 设$\dot{p}$ 是n维仿射空间$(\mathbb{A}, V)$的一个固定的点, 而U是V 的向量子空间. 那么, 集合

$$\Pi=\dot{p}+U=\{\dot{p}+\mathbf{u}|\mathbf{u}\in U\}$$

被称为是$\mathbb{A}$ 的一个$m=\dim U$ 维的**平面**(或者**仿射子空间**). 说Π是经过点$\dot{p}$ 的在向量子空间U方向上的仿射子空间. 当$m=0$时, 自然地, 就是一个**点**, 当$m=1$时, 是一条**直线**, 当$m=n-1$时, 它就是一个**超平面**(与向量空间中的术语、规矩完全对应). 还要说, U是平面Π的**方向子空间**.

注意, 如果$\dot{q}=\dot{p}+\mathbf{u}$, $\dot{r}=\dot{p}+\mathbf{v}$,$\mathbf{u}, \mathbf{v}\in U$, 那么

$$\dot{q}+(\mathbf{v}-\mathbf{u})=\dot{p}+\mathbf{u}+(\mathbf{v}-\mathbf{u})=\dot{p}+\mathbf{v}=\dot{r}.$$

从而$\overrightarrow{qr}=\mathbf{v}-\mathbf{u}$. 又因为$\mathbf{v}-\mathbf{u}\in U$, 所以$\overrightarrow{qr}\in U$, 可见

$$\dot{q},\dot{r}\in\Pi\Rightarrow\overrightarrow{qr}\in U. \tag{4}$$

其次,

$$\dot{s},\dot{q},\dot{r}\in\Pi\Rightarrow\dot{s}+\overrightarrow{qr}\in\Pi, \tag{5}$$

这是因为$\dot{s}=\dot{p}+\mathbf{w}$,$\mathbf{w}\in U$且$\overrightarrow{qr}\in U$, 所以, $\dot{s}+\overrightarrow{qr}=\dot{p}+(\mathbf{w}+\overrightarrow{qr})$, 其中$\mathbf{w}+\overrightarrow{qr}\in U$.

反之, 具有性质(4)和(5)的子集$\Pi\subset\mathbb{A}$, 显然是在我们所定义的意义下的一个平面.

也就是说, 方向子空间$U\subset V$ 作为所有形如$\overrightarrow{qr}$, $\dot{q}$, $\dot{r}\in\Pi$ 的向量的集合, 由平面Π唯一确定. 在定义Π时指明的点$\dot{p}$ 可以用另外的任意一点$\dot{q}\in\Pi$ 来代替. 事实上, $\dot{q}=\dot{p}+\mathbf{u}$, $\mathbf{u}\in U$, 所以

$$\dot{q}+U=(\dot{p}+\mathbf{u})+U=\dot{p}+(\mathbf{u}+U)=\dot{p}+U.$$

由Π的已知性质可以直接得到

定理3 *在一个仿射空间中的平面$\Pi = \dot{p} + U$ 本身也是个仿射空间, 它与向量空间U相伴.*

证明 事实上, 用U替换V , $\mathbb{A}$ 满足的公理i), ii)在Π上照样得到满足. 其次, 正如我们已经知道的, 对任意两点$\dot{q}$, $\dot{r} \in \Pi$, 向量$\mathbf{w} = \overrightarrow{qr}$ 属于U而且$\dot{r} = \dot{q} + \mathbf{w}$, 同时$\mathbf{w}$在V 中是被唯一确定了的, 这就意味着, 它在U中. □

我们还要得出一些关于仿射空间($\mathbb{A}$, V)的子空间的有用的事实. 进一步假定, 在这里, 暂时地, 基础域$\mathfrak{K}$的特征不等于2. 对应于一般定义, 在$r>0$维的平面Π上, 至少有两个不同的点$\dot{p}$, $\dot{q}$. 当$r = 1$时(Π是条直线), 我们有

$$\Pi = \{\dot{p} + \lambda\overrightarrow{pq} | \lambda \in \mathfrak{K}\}. \tag{6}$$

定理4 *子集$\Pi \subset \mathbb{A}$ 是个子空间(平面), 当且仅当, 它能整个包含通过自己上面两个不同点的直线*(char $\mathfrak{K} \neq 2$).

证明 先设Π是个平面. 那么, $\Pi = \dot{p} + U$, $\dot{p} \in \mathbb{A}$, $U \subset V$ 是个向量子空间. 如果$\dot{q}_1$, $\dot{q}_2 \in \Pi$, 那么, 按照(6), 经过$\dot{q}_1$, $\dot{q}_2$ 的直线上的点必然形如

$$\dot{q}_1 + \lambda\overrightarrow{q_1q_2} = \dot{p} + \overrightarrow{pq_1} + \lambda\overrightarrow{q_1q_2}.$$

如果$\dot{q}_1 = \dot{p} + \mathbf{u}_1$, $\dot{q}_2 = \dot{p} + \mathbf{u}_2$, 那么, $\mathbf{u}_1$, $\mathbf{u}_2 \in U$, $\overrightarrow{q_1q_2} = \mathbf{u}_2 - \mathbf{u}_1$, $\overrightarrow{pq_1} = \mathbf{u}_1$, 这就意味着

$$\dot{q}_1 + \lambda\overrightarrow{q_1q_2} = \dot{p} + \mathbf{u}_1 + \lambda(\mathbf{u}_2 - \mathbf{u}_1) \in \dot{p} + U = \Pi.$$

反过来, 设$\dot{p} \in \Pi$, $U = \{\overrightarrow{pq} | \dot{q} \in \Pi\}$. 需要证明, U是V中的向量子空间. 根据条件, 如果$\dot{q}_1$, $\dot{q}_2 \in \Pi$, $\overrightarrow{pq_1} = \mathbf{u}_1$,$pq_2 = \mathbf{u}_2$. 那么, 在$\mathbb{A}$ 的直线$\{\dot{q}_1 + \mu\overrightarrow{q_1q_2} | \mu \in \mathfrak{K}\}$上的点$\dot{p} + \mathbf{u}_1 + \lambda(\mathbf{u}_2 - \mathbf{u}_1)$, 对所有的$\lambda \in \mathfrak{K}$, 都必然在$\Pi$上. 换言之,

$$\mathbf{u}_1, \mathbf{u}_2 \in U \Rightarrow \mathbf{u}_1 + \lambda(\mathbf{u}_2 - \mathbf{u}_1) \in U.$$

此外, 因为$\dot{p} \in \Pi$, 所以$\mathbf{0} \in \Pi$. 当$\mathbf{u}_1 = \mathbf{0}$ 时, 我们得到蕴涵式$\mathbf{u}_2 \in U \Rightarrow \lambda\mathbf{u}_2 \in U$. 当$\lambda = \dfrac{1}{2}$ 时, 由$\mathbf{u}_1$, $\mathbf{u}_2 \in U$ 推出$\dfrac{1}{2}\mathbf{u}_1 + \dfrac{1}{2}\mathbf{u}_2 \in U$, 从而$\mathbf{u}_1 + \mathbf{u}_2 = 2\left(\dfrac{1}{2}\mathbf{u}_1 + \dfrac{1}{2}\mathbf{u}_2\right) \in U$, 可见$U$是$V$ 的一个向量子空间. □

推论 *如果Π' 和Π'' 都是仿射空间$\mathbb{A}$ 的平面, 那么, 它们的交集$\Pi = \Pi' \cap \Pi''$ 或者为空集, 或者也是$\mathbb{A}$ 的一个平面. 如果U' , U'' 和U分别是Π' , Π'' 和Π所对应的V的向量子空间, 那么, $U = U' \cap U''$.*

证明 如果Π只含有一个点, 那么, 命题成立(U是零子空间). 现在设Π至少含两个不同的点$\dot{q}_1$, $\dot{q}_2$. 那么, 按照定理4, 通过$\dot{q}_1$ 和$\dot{q}_2$ 的直线将整个地被包含在Π'中, 同样也在Π'' 中. 从而, 这条直线就整个地被包含在$\Pi = \Pi' \cap \Pi''$ 中. 再对照定理4, 我们就推出结论, Π是个平面. 话又说回来了, 这本来就是直接而且显然的: 如果$\dot{p} \in \Pi' \cap \Pi''$, 那么, $\Pi' = \dot{p} + U'$, $\Pi'' = \dot{p} + U''$. 此时, $\dot{q} \in \Pi' \cap \Pi'' \Rightarrow \dot{q} = \dot{p} + \mathbf{u}' = \dot{p} + \mathbf{u}''$,

其中$\mathbf{u}' = \mathbf{u}'' \in \Pi' \cap \Pi''$. 我们看到, Π是由所有形如$\dot{p} + \mathbf{u}$, $\mathbf{u} \in U' \cap U''$ 的点组成. 可见, 它是个与$U = U' \cap U''$ 相伴的平面. □

定义5 称具有同一个子空间U为方向子空间的任意两个平面是**平行的**.

显然, 两个平行的平面$\dot{p} + U$, $\dot{q} + U$ 重合, 当且仅当, $\overrightarrow{pq} \in U$. 在任何情形, 都有

$$\dot{q} + U = t_{\overrightarrow{pq}}(\dot{p} + U),$$

也就是说, 平行的平面可以由相互平移得到.

现在, 我们把定理2前面的那段说明弄得更仔细些.

定义6 称仿射空间$\mathbb{A}$ 的点$\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_n$ 处于**普通位置**(或者是**仿射无关的**), 如果它们不都属于任何一个$(m-1)$维的平面.

只有在$m \leqslant n = \dim \mathbb{A}$ 时, 点$\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_m$ 才可能有处于普通位置的性质, 等价于向量$\overrightarrow{p_0p_1}, \overrightarrow{p_0p_2}, \cdots, \overrightarrow{p_0p_m}$ 线性无关的条件, 或者又等价于, 对任意的另外一个指标i,

$$\overrightarrow{p_ip_0}, \cdots, \overrightarrow{p_ip_{i-1}}, \overrightarrow{p_ip_{i+1}}, \cdots, \overrightarrow{p_ip_m}$$

线性无关, 因为$\overrightarrow{p_ip_j} = \overrightarrow{p_0p_j} - \overrightarrow{p_0p_i}$. 把线性包络$\langle\overrightarrow{p_0p_1}, \cdots, \overrightarrow{p_0p_m}\rangle$ 与U联系起来, 就可以推出这样的事实, 即通过处于普通位置的点$\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_m$ 能够引出, 而且唯一地引出一个m维平面$\dot{p}_0 + U$.

在$\mathbb{A}$ 的任意的点集$\mathcal{M}$的情形, 所有由$\dot{p}_0 \in \mathcal{M}$ 出发且以同样也属于$\mathcal{M}$ 的点$\dot{p}$ 为终点的向量构成线性包络U的维数, 等于空间$\langle\overrightarrow{p_0p} | p \in \mathcal{M}\rangle$ 的维数, 与点$\dot{p}_0$ 的选择无关. 平面$\Pi := A(\mathcal{M}) := \dot{p}_0 + U$ 可以看成是所有的包含$\mathcal{M}$ 的平面的交集.

定义7 称平面$\Pi = A(\mathcal{M})$是集合$\mathcal{M}$ 的**仿射包络**.

特别地, 当$\mathcal{M} = \{\Pi', \Pi''\}$ 时就可以讨论关于两个平面$\Pi', \Pi'' \subset \mathbb{A}$ 的仿射包络$A(\Pi', \Pi'')$ 问题. 容易看出, 由$\mathcal{M}$ 确定的仿射包络$A(\mathcal{M})$是唯一的; $A(\mathcal{M})$是包含Π' 和Π'' 的最小的平面.

例2 $\Pi' = \{\dot{p}\}$ 是个只含2维平面$(\mathbb{A}, \mathbb{R}^2)$中一个点$\dot{p}$ 的零维的平面, $\Pi'' = \{\dot{q} + \lambda\overrightarrow{qr} | \lambda \in \mathbb{R}\}$ 是$(\mathbb{A}, \mathbb{R}^2)$的一条直线. 从而, 自然地, 只要$\dot{p} \in \Pi''$, 则$A(\Pi', \Pi'') = \Pi''$, 如果$\dot{p} \notin \Pi''$, 则$A(\Pi', \Pi'') = (\mathbb{A}, \mathbb{R}^2)$.

5. 重心坐标 在第3目的说明中已经产生了在定义3中用向量$\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$代替处于普通位置的点$\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_n$ 的思想. 任意点$\dot{p} \in \mathbb{A}$ 的坐标可由记法$\dot{p} = \dot{p}_0 + \sum_{i=1}^{n} x_i(\dot{p}_i - \dot{p}_0)$ 来决定. 这个表达式从形式上可以变成

$$\dot{p} = \left(1 - \sum_{i=1}^{n} x_i\right)\dot{p}_0 + \sum_{i=1}^{n} x_i\dot{p}_i$$

的样子, 其中各个分开的被加项没有任何意义. 更准确些, 对任意点$\dot{q}$, $\dot{r} \in \mathbb{A}, \lambda \in \mathfrak{K}$, 和$\dot{q} + \dot{r}$ 以及表达式$\lambda\dot{q}$ 没有任何几何学上的解释, 只有当$\mathbb{A} = V$ 的时候可以除外. 然而, 有意义的是

定义8[#] 设$\dot{p}_0,\ \dot{p}_1,\ \cdots,\ \dot{p}_m$是仿射空间$\mathbb{A}$的任意的点. 我们把任意一组满足条件$\sum_{i=0}^{m}\alpha_i = 1$的纯量$\alpha_0,\ \alpha_1,\ \cdots,\ \alpha_m \in \mathfrak{K}$与一个形式上的和$\sum_{i=0}^{m}\alpha_i\dot{p}_i$放在一起, 令

$$\sum_{i=0}^{m}\alpha_i\dot{p}_i = \dot{p} + \sum_{i=0}^{m}\alpha_i\left(\dot{p}_i - \dot{p}\right),$$

其中$\dot{p}$是$\mathbb{A}$中任意一个点. 说$\sum_{i=0}^{m}\alpha_i\dot{p}_i$是点$\dot{p}_0,\ \dot{p}_1,\ \cdots,\ \dot{p}_m$的带有系数$\alpha_0,\ \alpha_1,\ \cdots,\ \alpha_m$的**重心组合**.

此定义是合理的, 因为成立如下的命题

命题1 表达式

$$\sum_{i=0}^{m}\alpha_i\dot{p}_i := \dot{p} + \sum_{i=0}^{m}\alpha_i\left(\dot{p}_i - \dot{p}\right), \quad \sum_{i=0}^{m}\alpha_i = 1,$$

与点$\dot{p}$的选择无关.

证明 实际上, 用点$\dot{q} = \dot{p} + \mathbf{v}$, $\mathbf{v} \in V$来代替$\dot{p}$, 我们就得到

$$\begin{aligned}&\dot{p} + \mathbf{v} + \sum_{i=0}^{m}\alpha_i\left(\dot{p}_i - \dot{p} - \mathbf{v}\right)\\ &\quad = \dot{p} + \mathbf{v} + \sum_{i=0}^{m}\alpha_i\left(\dot{p}_i - \dot{p}\right) - \left(\sum_{i=0}^{m}\alpha_i\right)\mathbf{v} = \dot{p} + \sum_{i=0}^{m}\alpha_i\left(\dot{p}_i - \dot{p}\right),\end{aligned}$$

这是因为$\left(1 - \sum_{i=0}^{m}\alpha_i\right)\mathbf{v} = \mathbf{0}$. □

例如, 可以谈一下所谓"点的半和"$\frac{1}{2}\dot{q} + \frac{1}{2}\dot{r} = \dot{q} + \frac{1}{2}\left(\dot{r} - \dot{q}\right)$, 但是, 绝不可以谈论所谓"点的三分之一和" $\frac{1}{3}\dot{q} + \frac{1}{3}\dot{r}$.

定义8 如果任何一点$\dot{p} \in \mathbb{A}$都可以唯一地表达成重心组合形式

$$\dot{p} = \sum_{i=0}^{n}x_i\dot{p}_i, \quad x_i \in \mathfrak{K}, \quad \sum_{i=0}^{n}x_i = 1,$$

那么, 称点组$\{\dot{p}_0,\ \dot{p}_1,\ \cdots,\ \dot{p}_n\}$是$\mathbb{A}$的一个**重心坐标系**, 而$x_0,\ \cdots,\ x_n$被称为点$\dot{p}$的**重心坐标**.

再把对于点$\dot{p}$的表达式写成$\dot{p} = \dot{p}_0 + \sum_{i=1}^{n}x_i\left(\dot{p}_i - \dot{p}_0\right)$的形式, 我们可以看出, 重心组合的唯一性等价于系$\{\dot{p}_0;\ \dot{p}_1 - \dot{p}_0,\ \cdots,\ \dot{p}_n - \dot{p}_0\}$处于普通位置, 而这组向量$(\dot{p}_1 - \dot{p}_0,\ \cdots,\ \dot{p}_n - \dot{p}_0)$就是$V$的一个基底. 点$\dot{p}_0 + \mathbf{x}$的重心坐标可以唯一地按照向量$\mathbf{x}$的坐标$x_1,\ \cdots,\ x_n$复原成$x_0 = 1 - \sum_{i=1}^{n}x_i$, $x_1,\ \cdots,\ x_n$.

以略有不同的方式加以推断, 我们设$\{\dot{o};\ \mathbf{e}_1,\ \cdots,\ \mathbf{e}_n\}$是$n$维仿射空间$\mathbb{A}$的任意一个坐标架, Π_m是通过处于普通位置的点$\dot{p}_0,\ \dot{p}_1,\ \cdots,\ \dot{p}_m$唯一确定的$m$维平面,

且$x_1^i, \cdots, x_n^i$, $0 \leqslant i \leqslant m$ 是它们的坐标. 那么, 任意点$\dot{p} \in \Pi_m$ 的坐标$x_1, \cdots, x_n$ 有唯一的方式表达成

$$x_j = x_j^0 + \lambda_1 \left(x_j^1 - x_j^0\right) + \cdots + \lambda_m \left(x_j^m - x_j^0\right), \quad j = 1, 2, \cdots, n. \tag{7}$$

这些方程式提供了一种可以称之为Π_m 的**参数表达式**的方法. 如果我们引进一个与$\lambda_1, \cdots, \lambda_m$ 有关系的参数λ_0 满足关系式$\lambda_0 + \lambda_1 + \cdots + \lambda_m = 1$, 就可以取得更进一步的对称性. 进而可将(7)改写成

$$x_j = \lambda_0 x_j^0 + \lambda_1 x_j^1 + \cdots + \lambda_m x_j^m, \quad j = 1, 2, \cdots, n \tag{7'}$$

的样子.

正如从下面的命题可以看到的, 重心组合是可以很好地与仿射映射一致的.

命题2 i) 设$f : \mathbb{A} \to \mathbb{A}'$ 是个仿射映射, 且$\dot{p}_0, \cdots, \dot{p}_m \in \mathbb{A}$. 那么

$$f\left(\sum_{i=0}^{m} x_i \dot{p}_i\right) = \sum_{i=0}^{m} x_i f\left(\dot{p}_i\right), \quad \sum_{i=0}^{m} x_i = 1.$$

ii) 设点$\dot{p}_0, \dot{p}_1, \cdots, \dot{p}_n$ 给出$\mathbb{A}$ 的一个重心坐标组合. 那么, 对任意点$\dot{q}_0, \dot{q}_1, \cdots, \dot{q}_n \in \mathbb{A}'$ 都存在唯一的一个仿射映射f使得

$$f(\dot{p}_i) = \dot{q}_i, \quad i = 0, 1, \cdots, n.$$

证明 对于断言i), 我们可以选择一个点$\dot{p} \in \mathbb{A}$, 就得到

$$\begin{aligned} f\left(\sum_{i=0}^{m} x_i \dot{p}_i\right) &= f\left(\dot{p} + \sum_{i=0}^{m} x_i \left(\dot{p}_i - \dot{p}\right)\right) \\ &= f(\dot{p}) + Df\left(\sum_{i=0}^{m} x_i \left(\dot{p}_i - \dot{p}\right)\right) = f(\dot{p}) + \sum_{i=0}^{m} x_i Df\left(\dot{p}_i - \dot{p}\right) \\ &= f(\dot{p}) + \sum_{i=0}^{m} x_i \left(f\left(\dot{p}_i\right) - f\left(\dot{p}\right)\right) = \sum_{i=0}^{m} x_i f\left(\dot{p}_i\right), \end{aligned}$$

从而证明了断言i).

因为$\mathbb{A}$ 中所有的点都可以唯一地表示成一个重心组合, 那么, 用公式

$$f\left(\sum x_i \dot{p}_i\right) = \sum x_i \dot{q}_i$$

定义一个集合论中的映射$f : \mathbb{A} \to \mathbb{A}'$. 根据断言i), 这是唯一可能的定义, 从而只需验证f是个仿射映射就行了. 事实上

$$\begin{aligned} f\left(\sum_{i=0}^{n} x_i \dot{p}_i\right) - f\left(\sum_{i=0}^{n} y_i \dot{p}_i\right) &= \sum_{i=0}^{n} x_i \dot{q}_i - \sum_{i=0}^{n} y_i \dot{q}_i \\ &= \dot{q}_0 + \sum_{i=1}^{n} x_i \left(\dot{q}_i - \dot{q}_0\right) - \left(\dot{q}_0 + \sum_{i=1}^{n} y_i \left(\dot{q}_i - \dot{q}_0\right)\right) \\ &= \sum_{i=1}^{n} \left(x_i - y_i\right)\left(\dot{q}_i - \dot{q}_0\right) = Df\left(\sum_{i=0}^{n} x_i \dot{p}_i - \sum_{i=0}^{n} y_i \dot{p}_i\right), \end{aligned}$$

其中$Df: V \to V'$ 是线性映射, 对所有的$i = 1, \cdots, n$, 它把$\dot{p}_i - \dot{p}_0$ 变成$\dot{q}_i - \dot{q}_0$. 它是存在的, 因为按假定, $(\dot{p}_1 - \dot{p}_0, \cdots, \dot{p}_n - \dot{p}_0)$ 是向量空间V 的基底. □

例3 在实的仿射空间$\mathbb{A}$ 中, 任意一个非退化的三角形的三个顶点都可以构成一个坐标架. 比方说, 如果(1,0,0), (0,1,0),(0,0,1)是三角形顶点的重心坐标, 那么, $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ 就是这个三角形的重力中心的重心坐标.

6. 仿射线性函数与线性方程组 设$(\mathbb{A}, V)$是域$\mathfrak{K}$上的一个仿射空间. 与定义2相对应, 称$f: \mathbb{A} \to \mathfrak{K}$是个**仿射线性函数**, 如果

$$f(\dot{p} + \mathbf{v}) = f(\dot{p}) + Df \cdot \mathbf{v} \qquad \forall \dot{p} \in \mathbb{A}, \mathbf{v} \in V,$$

其中$Df \in V^*$ 是V 上的一个线性函数. 按前面的说法, 称它是函数f的线性部分(或微商), 常量, 也就是纯量, 对应线性部分为零的仿射线性映射.

选取一个坐标系$\{\dot{o}; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$ 并用$x_1, \cdots, x_n$ 代表点$\dot{p} \in \mathbb{A}$ 的坐标, 我们把函数f的值表达成

$$f(\dot{p}) = f(\dot{o} + \overrightarrow{op}) = \sum_{i=1}^{n} \alpha_i x_i + \alpha_0, \tag{8}$$

其中$\alpha_0 = f(\dot{o})$, $\alpha_i = Df \cdot \mathbf{e}_i$, $\overrightarrow{op} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$.

反之, 如果函数$f: \mathbb{A} \to \mathfrak{K}$的值可以按照公式(8)计算出来, 且$\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n$, 那么, 遵照定理2, 有

$$\begin{aligned} f(\dot{p} + \mathbf{v}) &= \sum_{i=1}^{n} \alpha_i (x_i + v_i) + \alpha_0 \\ &= \left(\sum_i \alpha_i x_i + \alpha_0\right) + \sum_i \alpha_i v_i = f(\dot{p}) + Df \cdot \mathbf{v}, \end{aligned}$$

也就是说, f是个仿射线性函数.

说明 令$(\lambda f + \mu g)(\dot{p}) = \lambda f(\dot{p}) + \mu g(\dot{p})$, $D(\lambda f + \mu g) = \lambda Df + \mu Dg$, 我们发现, 所有$\mathbb{A} \to \mathfrak{K}$的仿射线性函数的集合$\mathcal{S}$ 就被赋予了向量空间结构: 如果$f, g \in \mathcal{S}$, 那么

$$\begin{aligned} (\lambda f + \mu g)(\dot{p} + \mathbf{v}) &= \lambda f(\dot{p} + \mathbf{v}) + \mu g(\dot{p} + \mathbf{v}) \\ &= \lambda \{f(\dot{p}) + Df \cdot \mathbf{v}\} + \mu \{g(\dot{p}) + Dg \cdot \mathbf{v}\} \\ &= (\lambda f + \mu g)(\dot{p}) + (\lambda Df + \mu Dg) \cdot \mathbf{v} \\ &= (\lambda f + \mu g)(\dot{p}) + (D(\lambda f + \mu g)) \cdot \mathbf{v}, \end{aligned}$$

也就是说, $\lambda f + \mu g \in \mathcal{S}$. 现在, 可以把$D$解释成$\mathcal{S} \to V^*$ 上的线性映射. 核Ker D就是由所有常数函数组成的$\mathcal{S}$ 中的直线$\mathcal{S}^0$.

我们再一次地回到线性方程组

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_n \end{aligned} \tag{9}$$

上来, 并可以写成

$$f_1(\dot{p})=0,\cdots,f_m(\dot{p})=0, \tag{9'}$$

其中f_i 是仿射线性函数:

$$f_i(\dot{p})=\sum_{i=1}^{n}a_{ij}x_j-b_i.$$

设方程组(9)是相容的, 而x_1^0, $\cdots$, x_n^0 是它的一个解. 取x_1^0, $\cdots$, x_n^0 作为在某个坐标架$\{\dot{o};\ \mathbf{e}_1,\ \cdots,\ \mathbf{e}_n\}$之下点$\dot{p}_0$ 的坐标(因此, $f_i\left(\dot{o}\right)=-b_i$). 为简便起见, 我们约定, 就说点$\dot{p}_0$ 本身是个**解**, 在已知事实的基础上, 可以推出这样的结论, 方程组(9)或(9′)的其他解必然形如$\dot{p}=\dot{p}_0+\mathbf{x}$, 其中$\mathbf{x}\in V$ 且满足线性方程组

$$Df_1\cdot\mathbf{x}=0,\ \cdots,\ Df_m\cdot\mathbf{x}=0. \tag{10}$$

在这里, Df_i 是函数f_i 的线性部分, $Df_i\cdot\mathbf{x}=\sum\limits_j a_{ij}x_j$. 方程组(10)的解, 众所周知, 构成一个子空间$U\subset V$, 维数为$n-r$, 其中r是系Df_1, $\cdots$, Df_m 的秩(见第1章§3的第5目). 这样一来, 系(3)的解的集合就是$n-r$维的平面$\Pi=\dot{p}_0+U$.

反过来, 任意平面$\Pi=\dot{p}_0+U\subset\mathbb{A}$ 可以给定一个线性方程组. 实际上, 根据第1章§3的定理4, $n-r$维的子空间$U\subset V$ 就是一个秩为r的形如(10)的线性方程组的解空间. 其次, 按照点$\dot{p}$ 属于Π的定义, 其充要条件是$\overrightarrow{p_0p}\in U$. 如果, x_1, $\cdots$, x_n 是点$\dot{p}$ 在选定的坐标架之下的坐标, 而x_1^0, $\cdots$, x_n^0 是点$\dot{p}_0$ 的坐标, 那么, $\overrightarrow{p_0p}=\sum\limits_j\left(x_j-x_j^0\right)\mathbf{e}_j$, 从而方程组可采用

$$\sum_{j=1}^{n}a_{ij}\left(x_j-x_j^0\right)=0,\qquad i=1,\cdots,m,$$

或者

$$\sum_{j=1}^{n}a_{ij}x_j=b_i,\qquad i=1,\cdots,m$$

的形式. 其中$b_i=\sum\limits_j x_j^0$. 这个系的秩, 依然等于r.

这就是说, 已经证明了

定理5 设$\mathbb{A}$ 是个n维仿射空间, $\mathbb{A}$ 中的所有坐标满足一个秩为r的相容的线性方程组的点构成$\mathbb{A}$ 的一个$n-r$维的平面Π. $\mathbb{A}$ 的任意一个平面都可以这样得到.

特别地, 超平面给出一个线性方程式

$$a_1x_1+a_2x_2+\cdots+a_nx_n=b.$$

把注意力停留在$r\leqslant m$情形的线性方程组(线性仿射函数f_i)上, 我们将$n-r$维的平面Π看成是r个超平面的交集. 在不相容的线性方程组的情形, 超平面的交集是空的.

7. 平面位置关系 设$(\mathbb{A}, V)$是个n维仿射空间. 把同一维数的平面间的平行概念加以推广, 我们引进

定义9 设$\Pi' = \dot{p} + U'$, $\Pi'' = \dot{q} + U''$ (U', U'' 分别是V 的k, l维的向量子空间)且$k \leqslant l$. 说平面Π' **平行于**Π'', 如果$U'' \subseteq U'$.

在$k = l$的情形, 就回到了从前的平行的概念上. 如果$\Pi'' \subseteq \Pi'$, 那么, 平行性条件就自动地得到了满足. 盘点一下已经在平面和线性方程组之间建立的联系, 我们可以断言下面定理是成立的.

定理6 对任意平面$\Pi \subset \mathbb{A}$ 和任意点$\dot{q} \in \mathbb{A}$ 都能找到一个, 而且是唯一的, 通过点$\dot{q}$ 且平行于Π的平面Π', 同时还有$\dim \Pi' = \dim \Pi$. 如果$\dot{q} \in \Pi$, 则$\Pi' = \Pi$. 如果$\dot{q} \notin \Pi$, 那么Π与Π' 不相交.

特别地, 由同一个坐标系, 用方程

$$a_1x_1 + \cdots + a_nx_n = b, \qquad a_1'x_1' + \cdots + a_n'x_n' = b',$$

给出的两个超平面Π和Π' 的平行性意味着就是变量系数简单地成比例, $a_i' = \lambda a_i$, $i = 1, 2, \cdots, n$. 加上重合条件$\Pi = \Pi'$, 就自然地意味着又加了一个限制, 对同一个$\lambda \in \mathfrak{R}$, 还有$b' = \lambda b$.

定义10 称既不平行又不相交的两个平面$\Pi, \Pi' \in \mathbb{A}$ 为**偏斜**的.

我们可以用某些数量估计来加深对平面间的相互位置的刻画. 首先, 如果平面$\Pi' = \dot{p} + U'$, $\Pi'' = \dot{q} + U''$ 是相交的, 而$\dot{o}$ 是它们的一个公共点, 那么, 正如我们所知, 仿射包络必形如:

$$\Pi := A(\Pi', \Pi'') = \dot{o} + W, \qquad W = U' + U''.$$

在此情形, 根据第1章§2的定理6, 必有

$$m := \dim \Pi = \dim W = k + l - i,$$

其中

$$k = \dim U', \quad l = \dim U'', \quad k \geqslant l, \quad i = \dim(U' \cap U''). \tag{11}$$

如果交集$\Pi' \cap \Pi''$ 是空的, 那么我们研究向量直线$V_1 = \{\lambda\overrightarrow{pq} | \lambda \in \mathfrak{R}\}$ 和子空间

$$W^{\circ} = U' + U'' + V_1 \subseteq V.$$

因为$\overrightarrow{pq} \notin U' + U''$ (见习题4.1.1), 所以,

$$m = \dim \Pi = \dim(U' + U'') + \dim V_1 = k + l - i + 1.$$

显然, 平面$\Pi^{\circ} = \dot{p} + W^{\circ}$ 包含$\Pi' = \dot{p} + U'$ 和$\Pi'' = \dot{q} + U'' = \dot{p} + \overrightarrow{pq} + U'' \subset \dot{p} + U'' + V_1$. 另一方面, 所有的包含$\Pi'$ 和Π'' 的平面必然都包含$\overrightarrow{pq}$和直线V_1, 从而就包含Π°. 换言

之, 当$\Pi'\cap\Pi''=\varnothing$ 的时候, 我们有$\Pi^{\circ}=A(\Pi',\Pi'')$. 这样一来,

$$m=\dim A(\Pi',\Pi'')=\begin{cases} k+l-i, & \text{当}\Pi'\cap\Pi''\neq\varnothing, \\ k+l-i+1, & \text{当}\Pi'\cap\Pi''=\varnothing. \end{cases} \tag{12}$$

由关系式(12)决定的整数的四元组

$$(i,k,l,m),\quad 0\geqslant i\geqslant k\geqslant l\geqslant m\geqslant n, \tag{13}$$

完全刻画了平面Π' 和Π'' 间的相互位置.

例4 设($\mathbb{A}$, $\mathbb{R}^3$)是实的3维仿射空间, Π' 和Π'' 是它里边的两条直线, 从而$n=3$, $k=l=1$. 这两条直线在$\mathbb{A}$ 中的相互位置的各种不同的情形是相当明显的而且可以由图6反映出来.

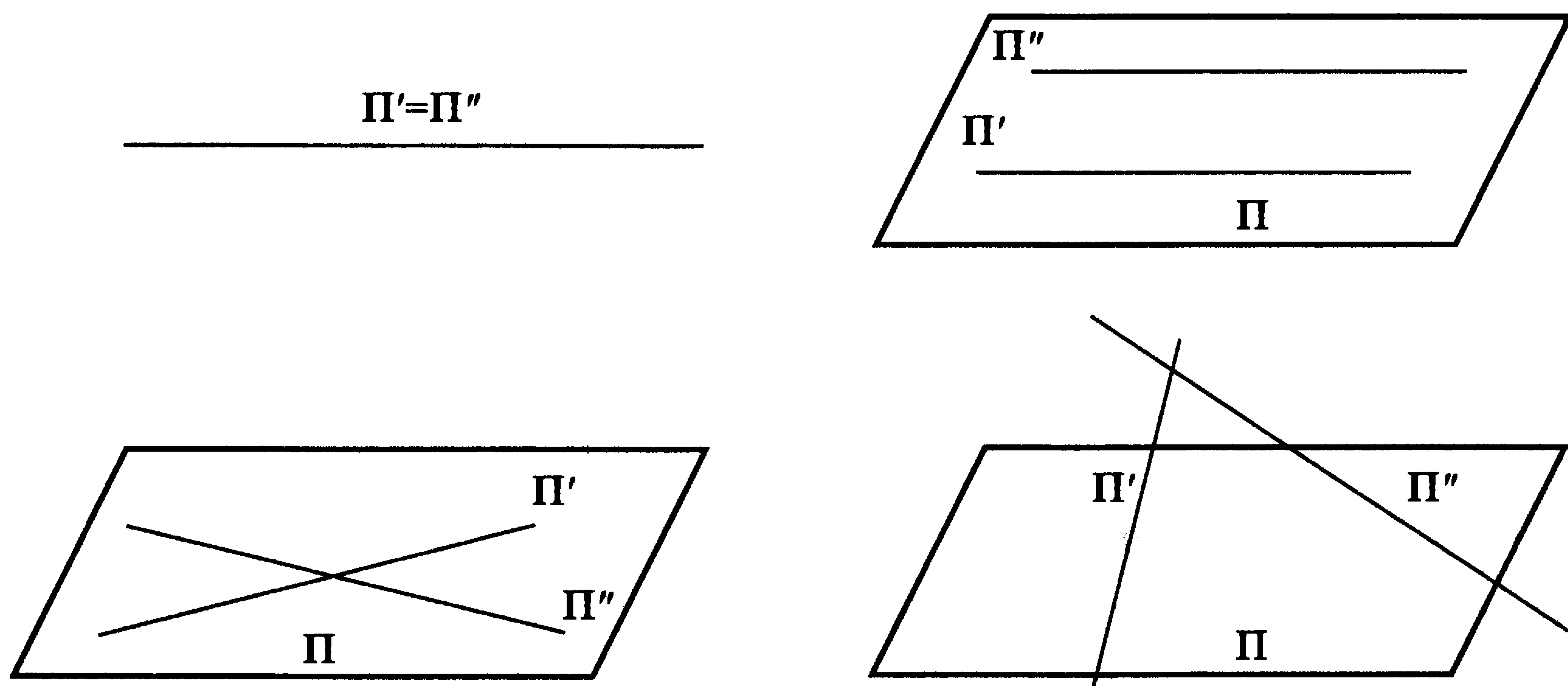


图6

可试着介绍一下在实的4维仿射空间中处于偏斜位置的一个1维平面和一个2维平面.

习 题

1. 试验证, 平面$\Pi'=\dot{p}+U'$, $\Pi''=\dot{q}+U''$ 是相交的, 即它们至少有一个共同点, 等价于$\overrightarrow{pq}\in U'+U''$ (见定理4的推论).

2. 设$A(\Pi_1,\cdots,\Pi_m)$ 是实的n维仿射空间$\mathbb{A}$ 中直线$\Pi_1,\cdots,\Pi_m$ 的仿射包络. 问, 在什么样的极小的m的情况下$A(\Pi_1,\cdots,\Pi_m)=\mathbb{A}$?

3. 设($\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_n$)是n维仿射空间$\mathbb{A}$ 的一个坐标架, 而($\dot{p}'_0$, $\dot{p}'_1$, $\cdots$, $\dot{p}'_n$) 是仿射空间$\mathbb{A}'$ 的一组$n+1$个点. 证明, 存在唯一的一个仿射映射$f:\mathbb{A}\to\mathbb{A}'$ 使得$f(\dot{p}_i)=\dot{p}'_i$, $i=0,1,\cdots,n$.

4. 证明, 点$\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_n$ 的有限多个重心组合的重心组合仍然是个重心组合.

5. 设$\mathbb{A}$ 是个n维仿射空间. 证明, 保持$\mathbb{A}$ 中点的重心组合不变的映射$f:\mathbb{A}\to\mathbb{A}$ 必然是个仿射变换(命题2中断言i)的逆命题).

6. 利用仿射变换的性质, 证明已知的关于三角形的三条中线必相交于一点的定理.

§2 欧几里得(点)空间

1. 欧几里得度量 为了紧密地贴近3维物理空间, 我们引进下列的

定义1 称仿射空间($\mathbb{E}$, V)是个**欧几里得(点)空间**, 如果V 是个欧几里得向量空间.

"点"这个词我们通常将其省略掉, 因为前面只研究过欧几里得向量空间, 从而不致引起任何混乱. 也就是说, 欧几里得空间是个三元组($\mathbb{E}$, V, ρ), 其中$\rho(*|*)$是$\mathbb{E}$ 中点与点之间的距离函数, 即

$$\rho\,(\dot{p},\dot{q}) := \|\overrightarrow{pq}\| = \sqrt{(\overrightarrow{pq}|\overrightarrow{pq})}. \tag{1}$$

这里, $(\mathbf{u}|\mathbf{v}\,)$是个由$V$ 上纯量乘积给出的正定型.

我们已经知道的(见第3章§3的第5目)度量空间中距离函数的性质可以分列如下:

i) $\rho\,(\dot{p},\dot{q}) = \rho\,(\dot{q},\dot{p})$;

ii) $\rho\,(\dot{p},\dot{q}) = 0 \Leftrightarrow \dot{p} = \dot{q}$;

iii) $\rho\,(\dot{p},\dot{q}) + \rho\,(\dot{q},\dot{r}) \geqslant \rho\,(\dot{p},\dot{r})$ (三角不等式).

往后, 我们将用符号$\Pi_{\dot{p},\dot{q}}$ 代表经过两个不同点$\dot{p}$, $\dot{q}$ 的直线.

定义2 把向量$\overrightarrow{pq}$和$\overrightarrow{rs}$之间的夹角φ

$$\cos\varphi = \frac{(\overrightarrow{pq}|\overrightarrow{rs})}{\|\overrightarrow{pq}\|\cdot\|\overrightarrow{rs}\|}$$

称为**直线$\Pi_{\dot{p},\dot{q}}$ 和$\Pi_{\dot{r},\dot{s}}$ 之间的夹角**.

定义3 在欧几里得空间($\mathbb{E}$, V)中, 把坐标系$\{\dot{o};\ \mathbf{e}_1,\ \cdots,\ \mathbf{e}_n\}$ 称为**直角坐标系**, 如果$\mathbf{e}_1,\ \cdots,\ \mathbf{e}_n$ 是欧几里得向量空间V的一个标准正交基底: $(\mathbf{e}_i|\mathbf{e}_j) = \delta_{ij},\ i,j = 1,2,\cdots,n$.

设, 在一个直角坐标系之下, $\mathbb{E}$ 中的点$\dot{p}$ 和$\dot{q}$ 的坐标分别是$x_1,\ x_2,\ \cdots,\ x_n$ 和$y_1,\ y_2,\ \cdots,\ y_n$. 那么, 向量$\overrightarrow{pq}$的坐标就是$y_1 - x_1,\ \cdots,\ y_n - x_n$. 所以, 对应于定义等式(1),

$$\rho\,(\dot{p},\dot{q}) = \sqrt{(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2} \tag{2}$$

就是一个可以按照它来测量出点与点之间距离的通用公式.

定理1 任意两个有限的维数相同的欧几里得(点)空间$\mathbb{E}$ 和$\mathbb{E}'$ 必然是同构的. 这意味着, 存在一个仿射空间的同构映射$f:\mathbb{E}\to\mathbb{E}'$, 它保持点之间的距离:

$$\rho\,(\dot{p},\dot{q}) = \rho'\,(f\,(\dot{p})\,,f\,(\dot{q})) \tag{3}$$

(ρ' 是$\mathbb{E}'$ 上的距离函数).

证明 在$\mathbb{E}$ 中选取一个直角坐标系$\{\dot{o};\ \mathbf{e}_1,\ \cdots,\ \mathbf{e}_n\}$, 在$\mathbb{E}'$ 中选取一个直角坐标系$\{\dot{o}';\ \mathbf{e}_1',\ \cdots,\ \mathbf{e}_n'\}$. 建立映射$f:\mathbb{E}\to\mathbb{E}'$, 使得

$$f\,(\dot{o}) = \dot{o}',\quad \mathcal{F}\,(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1\mathbf{e}_1' + \cdots + x_n\mathbf{e}_n'. \tag{4}$$

因为, 显然, 线性映射$\mathcal{F}$ 是个双射, 所以, 在§1定理1的证明中引进的验证可以表明, f是仿射空间$\mathbb{E}$ 和$\mathbb{E}'$ 之间的与$Df = \mathcal{F}$ 相伴的同构映射.

此外, 点$\dot{p}' = f(\dot{p})$ 在坐标系$\{\dot{o}';\ \mathbf{e}_1', \cdots, \mathbf{e}_n'\}$ 中的坐标就是$\dot{p}$ 在$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_n\}$ 之下的坐标$x_1, \cdots, x_n$. 而且, 由于在$\mathbb{E}$ 和$\mathbb{E}'$ 中距离$\rho(\dot{p},\dot{q})$ 和$\rho'(\dot{p}',\dot{q}')$ 可以用相同的公式(2)(借助基底的选择)计算出来, 所以, 欧几里得空间同构的条件(3)同样得到了满足. □

我们引进一些新的概念.

定义4 称集合

$$\dot{p}\dot{q} = \{\dot{p} + \lambda\overrightarrow{pq} | 0 \leqslant \lambda \leqslant 1\}$$

是仿射空间中连接点$\dot{p}$ 和点$\dot{q}$ 的**线段**.

按定义, $\dot{p}\dot{q} = \dot{q}\dot{p}$. 满足条件$\overrightarrow{pr} = \overrightarrow{rq}$ 的点$\dot{r} \in \dot{p}\dot{q}$ 通常称之为**线段**$\dot{p}\dot{q}$ **的中间点**. 在欧几里得空间情形, 线段$\dot{p}\dot{q}$ 的长度可以理解为量

$$|\dot{p}\dot{q}| := \|\overrightarrow{pq}\| = \rho(\dot{p},\dot{q}) .$$

2. 点到平面的距离 设Π是n维欧几里得空间的一个m维的平面, 而$\dot{p}$ 是位于平面Π外的一个点. 再设$\dot{q}$ 是Π上的一个点.

定义5 如果对任意$\dot{r}$, $\dot{s} \in \Pi$ 都有$(\overrightarrow{pq}|\overrightarrow{rs}) = 0$, 就说直线$\Pi_{\dot{p},\dot{q}}$ **垂直**于平面Π, 并记成$\Pi_{\dot{p},\dot{q}} \perp \Pi$; 在这种情形下, 称量$\rho(\dot{p},\dot{q})$ 是**点**$\dot{p}$ **到平面**Π**的距离**(如果$\dot{p} \in \Pi$, 它就等于零), 而$\dot{p}$, $\dot{q}$ 之间的线段$\dot{p}\dot{q}$ 就称为点$\dot{p}$ 在Π上的**垂直线**, 记成$\dot{p}\dot{q} \perp \Pi$.

垂直线的长度是由点$\dot{p}$ 到平面Π的最短的距离, 即, 对任意不同于$\dot{q}$ 的点$\dot{r} \in \Pi$, 都有$\rho(\dot{p},\dot{r}) > \rho(\dot{p},\dot{q})$. 事实上, 正如图7上可以看到的, $\overrightarrow{pr} = \overrightarrow{pq} + \overrightarrow{qr}$ 是两个正交的向量之和, 因此, 只要$\dot{r} \neq \dot{q}$ 就有

$$\rho(\dot{p},\dot{r})^2 = (\overrightarrow{pr}|\overrightarrow{pr}) = (\overrightarrow{pq}|\overrightarrow{pq}) + (\overrightarrow{qr}|\overrightarrow{qr}) = \rho(\dot{p},\dot{q})^2 + \rho(\dot{q},\dot{r})^2 > \rho(\dot{p},\dot{q})^2$$

(函数ρ的性质ii)).

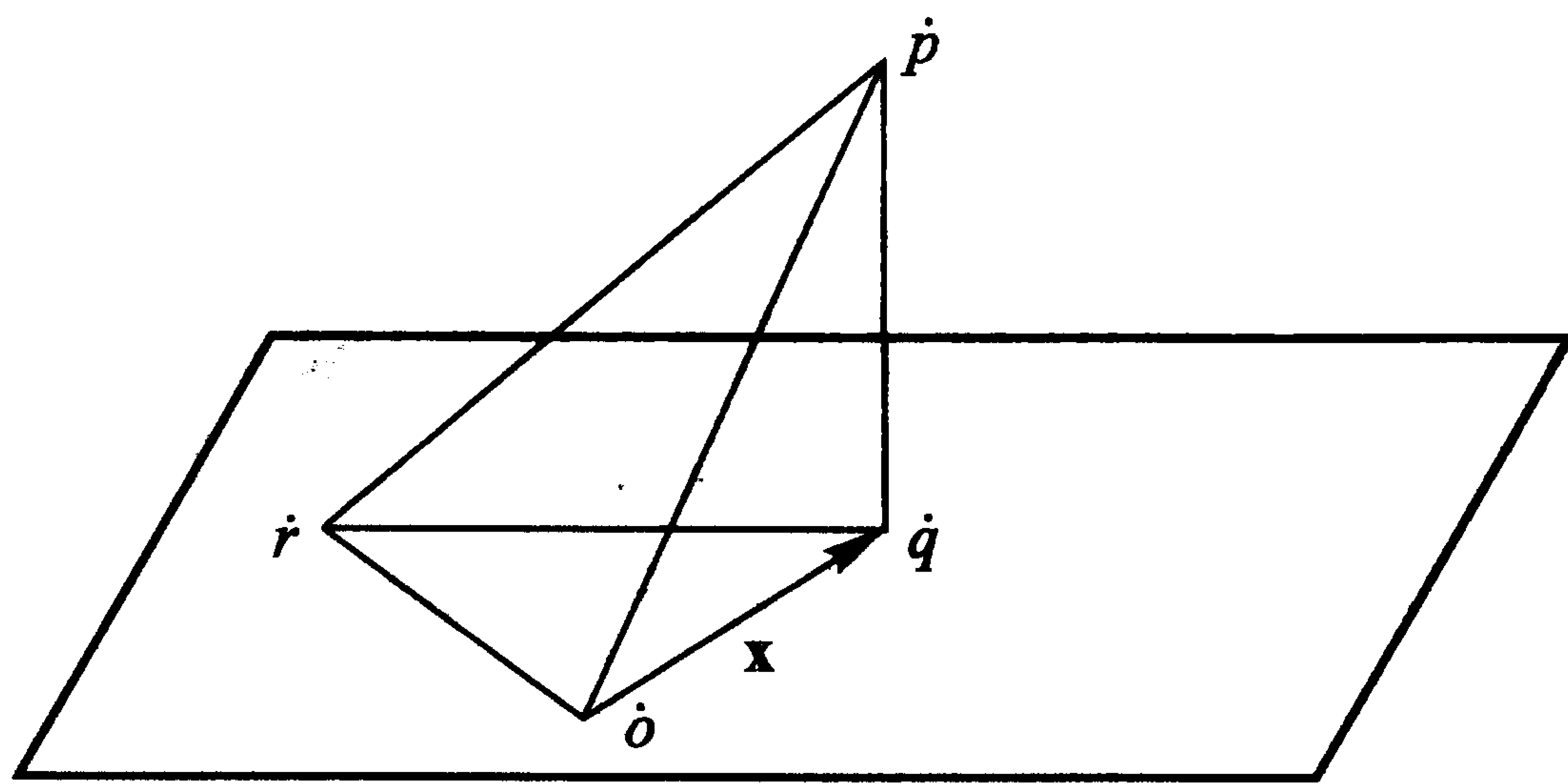


图7

设$\Pi = \dot{o} + U$. 由条件$\dot{p}\dot{q} \perp \Pi$, 把Π上点$\dot{q}$ 写成$\dot{q} = \dot{o} + \mathbf{x}$. 因为$V = U + U^{\perp}$ (第3章, §1的定理7), 而$\overrightarrow{op} = \mathbf{x} + \overrightarrow{qp}$, 其中$\mathbf{x} \in U$, 所以, 向量$\overrightarrow{op}$的分量$\overrightarrow{qp} \in U^{\perp}$ 存在并唯一确定.

为了具体地找到由$\dot{p}$ 到Π的垂直线, 我们在$\mathbb{E}$ 中选取直角坐标系

$$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_m, \cdots, \mathbf{e}_n\}, \tag{5}$$

其中$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 构成向量空间U的基底. 可以认为$\mathbf{v} = \overrightarrow{op}$ 是给定的. 计算向量的坐标

$$\mathbf{x} = \overrightarrow{op} + \overrightarrow{pq} = x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m,$$

我们就可以找到$h = \|\overrightarrow{pq}\|$. 注意到

$$\dot{p}\dot{q} \perp \Pi \Leftrightarrow (\overrightarrow{pq}|U) = 0 \Leftrightarrow (\overrightarrow{pq}|\mathbf{e}_i) = 0, \quad i = 1, \cdots, m.$$

从而, 有

$$(\mathbf{x} - \mathbf{v}|\mathbf{e}_i) = 0, \quad i = 1, \cdots, m, \tag{6}$$

进而得到$x_i = (\mathbf{v}|\mathbf{e}_i)$, $\quad i = 1, \cdots, m$.

如果坐标系(5)并不是一个直角坐标系, 那么, 条件(6)可表达成m个方程的线性方程组

$$(\mathbf{e}_1|\mathbf{e}_i)\, x_1 + (\mathbf{e}_2|\mathbf{e}_i)\, x_2 + \cdots + (\mathbf{e}_m|\mathbf{e}_i)\, x_m = (\mathbf{v}|\mathbf{e}_i), \quad i = 1, \cdots, m, \tag{7}$$

这个方程组已被证明有唯一解. m个未知元的线性方程组有唯一解, 当且仅当, 它的行列式不等于零. 方程组(7)的行列式是

$$G(\mathbf{e}_1, \cdots, \mathbf{e}_m) = \begin{vmatrix} (\mathbf{e}_1|\mathbf{e}_1) & \cdots & (\mathbf{e}_1|\mathbf{e}_m) \\ \cdots & \cdots & \cdots \\ (\mathbf{e}_m|\mathbf{e}_1) & \cdots & (\mathbf{e}_m|\mathbf{e}_m) \end{vmatrix}. \tag{8}$$

这样一来, 被称为是向量$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 的行列式$G(\mathbf{e}_1, \cdots, \mathbf{e}_m)$ 就不等于零. 我们又一次地得到了第3章§5的定理1, 实际上是重复了整个的推演.

把我们的有关于垂直性的知识归纳一下.

定理2 在欧几里得空间$\mathbb{E}$ 的平面$\Pi = \dot{o} + U$ 外面的任意一点$\dot{p}$ 都可以引一条垂直线$\dot{p}\dot{q}$. 它的长度$|\dot{p}\dot{q}|$ 是由$\dot{p}$ 到Π的最短距离. 如果$\dot{o}$ 是我们在Π上任意选取的一个点, 那么, $\overrightarrow{pq} = \mathbf{x} - \overrightarrow{op}$, 其中$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_m\mathbf{e}_m$ 是U的一个向量, 它在空间$\mathbb{E}$ 的任意一个坐标系$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_m, \cdots, \mathbf{e}_n\}$ 之下的坐标可以作为方程组的解按照克拉默公式计算出来

$$x_i = \frac{1}{G(\mathbf{e}_1, \cdots, \mathbf{e}_m)} \begin{vmatrix} (\mathbf{e}_1|\mathbf{e}_1) & \cdots & (\mathbf{e}_1|\mathbf{v}) & \cdots & (\mathbf{e}_1|\mathbf{e}_m) \\ (\mathbf{e}_2|\mathbf{e}_1) & \cdots & (\mathbf{e}_2|\mathbf{v}) & \cdots & (\mathbf{e}_2|\mathbf{e}_m) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (\mathbf{e}_m|\mathbf{e}_1) & \cdots & \underset{i}{(\mathbf{e}_m|\mathbf{v})} & \cdots & (\mathbf{e}_m|\mathbf{e}_m) \end{vmatrix}, \quad \mathbf{v} = \overrightarrow{op}.$$

如果$(\mathbf{e}_1, \cdots, \mathbf{e}_m)$是标准正交基底, 那么, $x_i = (\mathbf{e}_i|\overrightarrow{op})$.

3. 平面间的距离 设Π和Π'是欧几里得空间$(\mathbb{E}, V, \rho)$的两个平面, $\Pi = \dot{p} + U$, $\Pi' = \dot{p}' + U'$. 因为$\dot{p}'$可以用平面Π'上的任意一点来代替, 所以, 不失一般性, 可以认为$\dot{p}\dot{p}' \perp \Pi'$. 也就是说, $\Pi_{\dot{p},\dot{p}'} \perp \Pi'$. 如果同时还有$\Pi_{\dot{p},\dot{p}'} \perp \Pi$, 那么, 线段$\dot{p}\dot{p}'$就是到Π和Π'的公垂线.

引理1 *如果线段$\dot{p}\dot{p}'$是到Π和Π'的公垂线, 那么, 不管怎样的点$\dot{q} \in \Pi$, $\dot{q}' \in \Pi'$都有*

$$\rho(\dot{p}, \dot{p}') \leqslant \rho(\dot{q}, \dot{q}'). \tag{9}$$

证明 设$\dot{q} = \dot{p} + \mathbf{u}$, $\dot{q}' = \dot{p}' + \mathbf{u}'$. 因为$\dot{p}' = \dot{p} + \overrightarrow{pp'}$, 所以, $\dot{q}' = \dot{p} + \overrightarrow{pp'} + \mathbf{u}'$ 而且

$$\dot{q}\dot{q}' = \overrightarrow{pp'} + \mathbf{u}' - \mathbf{u}.$$

根据条件$\left(\overrightarrow{pp'}|\mathbf{u}\right) = 0$ 且$\left(\overrightarrow{pp'}|\mathbf{u}'\right) = 0$, 从而$\left(\overrightarrow{pp'}|\mathbf{u}' - \mathbf{u}\right) = 0$, 而在这种情形, 据毕达哥拉斯定理, 我们有

$$\left\|\overrightarrow{qq'}\right\|^2 = \|\mathbf{u}' - \mathbf{u}\|^2 + \left\|\overrightarrow{pp'}\right\|^2,$$

从而可得不等式(9). □

引理2 *在$(\mathbb{E}, V, \rho)$中任意两个平面Π, Π'间必定有一公垂线.*

证明 设$\Pi = \dot{q} + U$, $\Pi' = \dot{q}' + U'$. 选取点$\dot{p} = \dot{q} - \mathbf{u}$, $\dot{p}' = \dot{q}' - \mathbf{u}'$ 使得向量$\overrightarrow{pp'}$正交于U 和U' . 显然, $\overrightarrow{pp'} = \mathbf{u} - \mathbf{u}' + \overrightarrow{qq'}$. 因为. $V = (U + U') \oplus (U + U')^{\perp}$, 所以, $\overrightarrow{qq'} = \mathbf{b} + \mathbf{c}$, 其中$\mathbf{b} \in U + U'$, $\mathbf{c} \in (U + U')^{\perp}$, 而且$\mathbf{b}$和$\mathbf{c}$都是唯一确定的. 此外, $\mathbf{b} = \mathbf{v} + \mathbf{v}'$, $\mathbf{v} \in U$, $\mathbf{v}' \in U'$. 我们就得到

$$\overrightarrow{pp'} = -\mathbf{u}' + \mathbf{v}' + \mathbf{v} + \mathbf{u} + \mathbf{c}.$$

如果我们取$\mathbf{u}' = \mathbf{v}'$, $\mathbf{u} = -\mathbf{v}$, 向量$\overrightarrow{pp'}$即为所求. 事实上, $\overrightarrow{pp'} = \mathbf{c} \in (U + U')^{\perp}$. □

由引理1和引理2差不多可以直接推出

定理3 *对于任意两个平面Π, $\Pi' \subset (\mathbb{E}, V, \rho)$, 都能找到点$\dot{p} \in \Pi$, $\dot{p}' \in \Pi'$, 它们满足不等式(9). 线段$\dot{p}\dot{p}'$是Π和Π'的公垂线. 这条公垂线能够被唯一确定, 当且仅当, $U \cap U' = 0$ (U 和U'是Π和Π'的方向子空间).*

证明 事实上, 如果$\dot{p}\dot{p}'$和$\dot{q}\dot{q}'$是它们的两条公垂线, 那么$\rho(\dot{p}, \dot{p}') = \rho(\dot{q}, \dot{q}')$, 于是按照引理2的证明应当有$\mathbf{u} = \mathbf{u}'$, 即$\dot{q}' = \dot{p}' + \mathbf{u}$, $\dot{q} = \dot{p} + \mathbf{u}$, $\mathbf{u} \in U \cap U'$. 这样一来, 公垂线与$U \cap U'$的向量相互单值对应, 唯一性当且仅当$U \cap U' = 0$ 时成立. 特别地, 当Π'是一个点的时候, 即$U' = 0$时, 事情就更简单了.

4. 格拉姆行列式与平行六面体的体积 第2目中解决关于垂直线的问题把我们顺便引入对格拉姆行列式$G(\mathbf{e}_1, \cdots, \mathbf{e}_m)$ 不等于零的追求, 这个行列式是按公式(8)计算出来的, 它要求向量$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 线性无关. 可以把$G(\mathbf{e}_1, \cdots, \mathbf{e}_m)$看成是平面$\Pi$的方向子空间$U$上的二次型$q(\mathbf{v}) = (\mathbf{v}|\mathbf{v})$ 的矩阵的最后一个主子式Δ_m. 按着西尔维斯特准则(第1章§4的定理8), $G(\mathbf{e}_1, \cdots, \mathbf{e}_m) = \Delta_m > 0$.

如果$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 是线性相关的, 比如, 设$\mathbf{e}_m = \alpha_1\mathbf{e}_1 + \cdots + \alpha_{m-1}\mathbf{e}_{m-1}$ (我们已经给出子空间U的基底), 那么,

$$(\mathbf{e}_m|\mathbf{e}_i) = \sum_{i=1}^{m-1} \alpha_i (\mathbf{e}_m|\mathbf{e}_i), \quad i = 1, \cdots, m,$$

从而, $G(\mathbf{e}_1, \cdots, \mathbf{e}_m)$ 的最后一行就是其余各行的一个线性组合. 这样一来, $G(\mathbf{e}_1, \cdots, \mathbf{e}_m) = 0$, 且有下面的

定理4　向量组$\mathbf{e}_1, \cdots, \mathbf{e}_m$ 的格拉姆行列式不等于零, 当且仅当, 该向量组线性无关. 不等式$G(\mathbf{e}_1, \cdots, \mathbf{e}_m) \geqslant 0$ 恒成立, 它的特殊情形, 当m=2时, 就是柯西–布尼亚科夫斯基不等式.

格拉姆行列式可以解释成为一个以$\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_m$ ($\dot{p}_i = \dot{o} + \mathbf{e}_i, \quad i = 1, \cdots, m$)为边的平行六面体$P(\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_m)$ 的体积的平方v_m^2 :

$$P(\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_m) = \{t_1\dot{o}\dot{p}_1 + \cdots + t_m\dot{o}\dot{p}_m | 0 \leqslant t_i \leqslant 1\}. \tag{10}$$

设$(f_1, \cdots, f_m)$ 是向量空间U的一个标准正交基底, $\mathbf{e}_1, \cdots, \mathbf{e}_m$ 是U的一组向量(可能是线性相关的), 令

$$\mathbf{e}_j = \sum_{i=1}^{m} a_{ij}\mathbf{f}_i.$$

按定义我们计算出

$$v_m = v_m(P(\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_m)) = \left|\det(a_{ij})_1^m\right|, \tag{11}$$

这里与[BA Ⅰ]第3章§1的第1目的说明完全对应. 因为基底$(\mathbf{f}_1, \cdots, \mathbf{f}_m)$ 是标准正交的, 所以

$$\begin{aligned}(v_m)^2 &= \begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix}^2 = \begin{vmatrix} a_{11} & \cdots & a_{m1} \\ \cdots & \cdots & \cdots \\ a_{1m} & \cdots & a_{mm} \end{vmatrix} \cdot \begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix} \\ &= \det(\widetilde{a}_{ij}), \\ \widetilde{a}_{ij} &= \sum_{k=1}^{m} a_{ki}a_{kj} = (\mathbf{e}_i|\mathbf{e}_j),\end{aligned}$$

进而, 有

$$(v_m)^2 = \begin{vmatrix} (\mathbf{e}_1|\mathbf{e}_1) & \cdots & (\mathbf{e}_1|\mathbf{e}_m) \\ \cdots & \cdots & \cdots \\ (\mathbf{e}_m|\mathbf{e}_1) & \cdots & (\mathbf{e}_m|\mathbf{e}_m) \end{vmatrix} = G(\mathbf{e}_1, \cdots, \mathbf{e}_m).$$

总之, 格拉姆行列式等于一个以$\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_m$ 为边的m维的平行六面体的体积的平方.

在教学参考书[2]中给出了关于欧几里得空间中图形体积的更加充分的阐述.

习 题

1. 求出由点$\dot{p}=(2,1,-3,4)$ 到平面

$$\Pi: 2x_1-4x_2-8x_3+13x_4+19=0,\ \ x_1+x_2-x_3+2x_4-1=0$$

的距离.

2. 求出平面

$$\Pi_1: x_1+x_3+x_4-2x_5-2=0,\ \ x_2+x_3-x_4-x_5-3=0,\ \ x_1-x_2+2x_3-x_5-3=0$$

和平面

$$\Pi_2: (1,-2,5,8,2)+\langle(0,1,2,1,2),(2,1,2,-1,1)\rangle$$

之间的距离.

3. 说明, (11)式中的平行六面体的体积v_m 可以按照公式

$$v_m=\|\overrightarrow{op_1}\cdot l_1\cdots l_{m-1}\|$$

计算出来. 其中l_k 是由点$\dot{p}_{k+1}$ 到仿射包络$A(\dot{o},\dot{p}_1,\cdots,\dot{p}_m)$ 的垂直线的长度.

§3 群与几何

1. 仿射群 我们首先看一个例子.

例1 按定义, 实的仿射直线$\mathbb{A}$ 与实数集合$\mathbb{R}$ 是重合的. 换句话来说, 点$\dot{x}\in\mathbb{R}$ 可以与实数$x\in\mathbb{R}$ 等同起来(我们将利用这种等同). 直线的几何, 在§1第2目的意义下, 可以用仿射自同构来刻画. 在这种给定的情形, 映射$\mathbf{\Phi}_{\alpha,\beta}:\mathbb{A}\to\mathbb{A}$ 按规则

$$\mathbf{\Phi}_{\alpha,\beta}: x\mapsto\alpha x+\beta,\quad \alpha\in\mathbb{R}^*,\beta\in\mathbb{R} \tag{1}$$

来定义.

如果需要, x就是点$\dot{x}$ 在某个坐标架$\{\dot{o},\ \mathbf{e}\}$ 之下的坐标, 即$\dot{x}=\dot{o}+x$, $\mathbf{\Phi}_{\alpha,\beta}(\dot{x})=(\dot{o}+\beta)+\alpha x=\dot{o}+(\alpha x+\beta)$. 用符号$A_1=\mathrm{Aff}(\mathbb{R})$ 代表所有形如(1)的仿射变换的集合. 因为, 任意两个变换$\mathbf{\Phi}_{\alpha,\beta}$, $\mathbf{\Phi}_{\sigma,\tau}\in A_1$, 复合

$$\mathbf{\Phi}_{\alpha,\beta}\cdot\mathbf{\Phi}_{\sigma,\tau}=\mathbf{\Phi}_{\alpha\sigma,\alpha\tau+\beta} \tag{2}$$

仍然属于A_1, 又因为$\mathcal{E}=\mathbf{\Phi}_{1,0}\in A_1$, 而$\mathbf{\Phi}_{\alpha^{-1},-\alpha^{-1}\beta}$ 刚好是$\mathbf{\Phi}_{\alpha,\beta}$ 的逆变换, 所以, 集合A_1 连同自然的乘法运算(2)就构成了一个群(见[BA Ⅰ]第4章§2), 它被称为是**1维的实的仿射群**. 由(2)可得A_1是个非交换群, 且映射

$$\pi: \mathbf{\Phi}_{\alpha,\beta}\mapsto\alpha$$

是$A_1 \to \mathbb{R}^*$ 的一个满同态, 其中$\mathbb{R}^*$ 是所有非零实数构成的乘法群. 明显地, $\mathrm{Ker}\,\pi = \{\Phi_{1,\beta} | \beta \in \mathbb{R}\}$ 是A_1的所有的平移构成的子解, 它同构于所有实数构成的加法群$\mathbb{R}^+ = \{\mathbb{R}, +\}$. 我们就有了一个通常所说的群的态射的正合列:

$$0 \to \mathbb{R}^+ \to A_1 \to \mathbb{R}^* \to 1.$$

现在, 设$(\mathbb{A}\ , V\ )$是域$\mathfrak{K}$上的一个n维仿射空间, $f : \mathbb{A} \to \mathbb{A}$ 是个双射的仿射变换(仿射自同构). 按一般定义(见§1)

$$f\,(\dot{p} + \mathbf{v}) = f(\dot{p}) + \mathcal{F}\mathbf{v},$$

其中$\mathcal{F}$ 是V 上的非退化的线性算子, 一般情形, 用Df表示之. 根据条件, $\det\,\mathcal{F} \neq 0$, 即线性算子$\mathcal{F}$ 有逆算子$\mathcal{F}^{-1}$, 它是仿射变换f^{-1}的线性部分:

$$f^{-1}\,(\dot{p} + \mathbf{v}) = f^{-1}\,(\dot{p}) + \mathcal{F}^{-1}\mathbf{v}.$$

用e代表单位仿射变换(或者恒等变换), 它的线性部分是$\mathcal{E} : \mathbf{v} \mapsto \mathbf{v}$, 我们看到, $f \cdot f^{-1} = f^{-1} \cdot f = e$.

设f和g是空间$\mathbb{A}$ 的两个仿射变换, 它们的复合

$$h = f \cdot g : \dot{p} \mapsto f\,(g\,(\dot{p}))$$

依然是个仿射变换, 而且它的线性部分$\mathcal{H} = \mathcal{F}\mathcal{G}$ ($\mathcal{F}$ 和$\mathcal{G}$ 分别是变换f和g的线性部分). 实际上,

$$\begin{aligned} h\,(\dot{p} + \mathbf{v}) &= f\,(g\,(\dot{p} + \mathbf{v})) = f\,(g\,(\dot{p}) + \mathcal{G}\mathbf{v}) \\ &= f\,(g\,(\dot{p})) + \mathcal{F}\,(\mathcal{G}\mathbf{v}) = (f \cdot g)\,(\dot{p}) + \mathcal{F}\mathcal{G}\mathbf{v} = h\,(\dot{p}) + \mathcal{H}\mathbf{v}. \end{aligned}$$

所有仿射自同构的集合$\mathrm{Aff}\,(\mathbb{A}) = A_n(\mathfrak{K})$上的乘法运算的结合性可由所有$\mathbb{A} \to \mathbb{A}$ 的变换的集合上的复合的结合律推出来(见[BA Ⅰ]). 这样一来, $\mathrm{Aff}\,(\mathbb{A})$就是个群, 称它是仿射空间$\mathbb{A}$ 上的n维**仿射群**. 我们就得到了下述定理的特殊部分.

定理1 域$\mathfrak{K}$ 上的n 维仿射空间$(\mathbb{A}\ , V\ )$的所有的仿射自同构的集合$A_n(\mathfrak{K})$ 构成一个群. 所有的保持固定点$\dot{o}$不动的仿射自同构构成一个同构于完全线性群$GL(V) = GL_n(\mathfrak{K})$ 的子群$A_n(\mathfrak{K})_{\dot{o}} \subset A_n(\mathfrak{K})$. 空间$\mathbb{A}$ 的所有平移构成的子群$T = \{t_{\mathbf{v}} | \mathbf{v} \in V\}$ 在群$A_n(\mathfrak{K})$中是正规的, 而且属于正合列

$$e \to T \xrightarrow{\varphi} A_n(\mathfrak{K}) \xrightarrow{D} GL_n(\mathfrak{K}) \to \overline{e}.$$

中满射映射D的核.

(正合性意味着$\mathrm{Im}\,\varphi = \mathrm{Ker}\ D$).

证明 看仿射自同构$f, g \in A_n(\mathfrak{K})_{\dot{o}}$. 因为$f(\dot{o} + \mathbf{x}) = \dot{o} + \mathcal{F}\,\mathbf{x}$和$g(\dot{o} + \mathbf{x}) = \dot{o} + \mathcal{G}\mathbf{x}$, 所以$(f \circ g)(\dot{o} + \mathbf{x}) = f\,(g\,(\dot{o} + \mathbf{x})) = f\,(\dot{o} + \mathcal{G}\mathbf{x}) = \dot{o} + \mathcal{F}\,\mathcal{G}\,\mathbf{x}$, 也就是$f \circ g \in A_n(\mathfrak{K})_{\dot{o}}$.

可以类似地证明, $f^{-1}\in A_n(\mathfrak{K})_{\dot{o}}$. 单位映射同样也在$A_n(\mathfrak{K})_{\dot{o}}$ 中. 这样一来, $A_n(\mathfrak{K})_{\dot{o}}$ 就是$A_n(\mathfrak{K})$的一个子群. 对应$D: f\mapsto Df=\mathcal{F}$, 对任意$f\in A_n(\mathfrak{K})_{\dot{o}}$ 显然是群$A_n(\mathfrak{K})_{\dot{o}}$ 到所有非退化线性算子所构成的群$GL(V)$ 上的同构.

前面我们还曾注意过, 平移构成的子解$T\subset A_n(\mathfrak{K})$同构于空间$V$ 的加法群. 设$t_{\mathbf{v}}$是个平移, f是任意一个仿射自同构, 它的线性部分是$\mathcal{F}$, 那么

$$\begin{aligned}(f^{-1}t_{\mathbf{v}}f)(\dot{p})&=(f^{-1}t_{\mathbf{v}})f(\dot{p})=f^{-1}(f(\dot{p})+\mathbf{v})\\&=f^{-1}(f(\dot{p}))+\mathcal{F}^{-1}\mathbf{v}=\dot{p}+\mathcal{F}^{-1}\mathbf{v}=t_{\mathcal{F}^{-1}\mathbf{v}}(\dot{p}).\end{aligned}$$

因为$\dot{p}$ 是任意一个点, 从而可以推出

$$f^{-1}t_{\mathbf{v}}f=t_{\mathcal{F}^{-1}\mathbf{v}}. \tag{3}$$

等式(3)表明, T是$A_n(\mathfrak{K})$的正规子群, 也就是某个同态映射的核. 我们来找出这个同态映射.

我们已经足够熟悉了的映射$D: f\to Df=\mathcal{F}$ 的一个子群为它的核, $\operatorname{Ker} D=\{f\in A_n(\mathfrak{K})|\mathcal{F}=\mathcal{E}\}$. 这就意味着

$$f\in\operatorname{Ker} D\Rightarrow f(\dot{p}+\mathbf{v})=f(\dot{p})+\mathbf{v}.$$

在这种情况下, 向量$\mathbf{u}=\overrightarrow{(p+\mathbf{v})f(p+\mathbf{v})}=\overrightarrow{(p+\mathbf{v})(f(p)+\mathbf{v})}=\overrightarrow{pf(p)}$ 与点$\dot{p}$ 没有关系, 而且$f(\dot{p}+\mathbf{v})=(\dot{p}+\mathbf{u})+\mathbf{v}=(\dot{p}+\mathbf{v})+\mathbf{u}$, 所以, $f=t_{\mathbf{u}}$ 是个用向量$\mathbf{u}$做的平移. 总而言之, $\operatorname{Ker} D=T$. 另一方面, 任选一个初始点$\dot{o}$ 并建立任意一个$\mathcal{F}\in GL(V)$, 映射$\dot{o}+\mathbf{v}\mapsto\dot{o}+\mathcal{F}\mathbf{v}$, 我们可以看出, $\operatorname{Im} D=GL(V)$, 所以, D是个双射映射. □

我们还可以证明如下的断言.

定理2 *任意一个仿射变换$f:\mathbb{E}\to\mathbb{E}$, 它的线性部分是$\mathcal{F}$, 都可以表示成$f=t_{\mathbf{a}}g$ 的样子, 其中$t_{\mathbf{a}}$ 是加向量$\mathbf{a}=\overrightarrow{of(o)}$ 的平移, g是一个在给定点$\dot{o}$ 处保持不动的仿射变换. 这个分解与点$\dot{o}$ 有关系. 如果用点$\dot{o}'$ 代替点$\dot{o}$, 那么, 需要取$\mathbf{a}'=\mathbf{a}+(\mathcal{F}-\mathcal{E})\overrightarrow{oo'}$来代替$\mathbf{a}$.*

证明 令$\mathbf{a}=\overrightarrow{of(o)}$ 且$g=t_{\mathbf{a}}^{-1}f$. 我们已经知道, g必然是个仿射变换. 于是, $g(\dot{o})=t_{\mathbf{a}}^{-1}\cdot f(\dot{o})=t_{-\mathbf{a}}\cdot f(\dot{o})=f(\dot{o})-\overrightarrow{of(o)}=\dot{o}$. 由此可见, g保持$\dot{o}$ 不变.

取另外一点$\dot{o}'$ 代替$\dot{o}$, 我们又得到$\mathbf{a}'=\overrightarrow{o'f(o')}$. 如果$\dot{o}'=\dot{o}+\mathbf{b}$, 那么, $f(\dot{o}')=f(\dot{o})+\mathcal{F}\mathbf{b}$, 或者, 同样地, $\dot{o}'+\overrightarrow{o'f(o')}=\dot{o}+\overrightarrow{of(o)}+\mathcal{F}\mathbf{b}$, 从而有

$$\mathbf{a}'=\mathcal{F}\mathbf{b}-\mathbf{b}+\mathbf{a}=\mathbf{a}+(\mathcal{F}-\mathcal{E})\overrightarrow{oo'}.$$ □

在一个固定的原点$\dot{o}\in\mathbb{A}$, 群$\operatorname{Aff}(\mathbb{A})$可以把自己表达成对$(\mathcal{F},\mathbf{v})$的集合$(GL(V), V)$, $(\mathcal{F},\mathbf{v})$ 的作用是

$$(\mathcal{F},\mathbf{v})(\dot{o}+\mathbf{x})=\dot{o}+\mathcal{F}\mathbf{x}+\mathbf{v} \tag{4}$$

它的复合运算规律是

$$(\mathcal{F}_1, \mathbf{v}_1) \cdot (\mathcal{F}_2, \mathbf{v}_2) = (\mathcal{F}_1\mathcal{F}_2, \mathbf{v}_1 + \mathcal{F}_1\mathbf{v}_2). \tag{5}$$

实际上, 如果$f_i = (\mathcal{F}_i, \mathbf{v}_i), i = 1, 2$, 那么

$$\begin{aligned}(f_1 \cdot f_2) \cdot (\dot{o} + \mathbf{x}) &= f_1\left(f_2(\dot{o} + \mathbf{x})\right) \\ &= f_1\left(\dot{o} + \mathcal{F}_2\mathbf{x} + \mathbf{v}_2\right) = \dot{o} + \mathcal{F}_1\left(\mathcal{F}_2\mathbf{x} + \mathbf{v}_2\right) + \mathbf{v}_1.\end{aligned}$$

从而可得(4)和(5).

现在, 在$\mathbb{A}$ 中任意选择一个坐标系$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_n\}$. 那么, 按定义, 点$\dot{p} = \dot{o} + \mathbf{x}$的坐标就是向量$\overrightarrow{op} = \mathbf{x} = \sum_i x_i\mathbf{e}_i$ 的坐标$x_1, \cdots, x_n$. 如果f是个仿射变换, $Df = \mathcal{F}$是它的线性部分, 那么,

$$f(\dot{p}) = f(\dot{o}) + \mathcal{F}\mathbf{x} = \dot{o} + \overrightarrow{\dot{o}f(\dot{o})} + \mathcal{F}\mathbf{x}.$$

用$y_1, \cdots, y_n$ 代表点$f(\dot{p})$ 的坐标, 同样设$\overrightarrow{\dot{o}f(\dot{o})} = \sum_i b_i\mathbf{e}_i$, 而$F = (f_{ij})$ 是线性算子$\mathcal{F}$ 的矩阵, 使得

$$(\mathcal{F}\mathbf{x})_i = \sum_{j=1}^{n} f_{ij}x_j.$$

综上所述, 我们得到

$$y_i = \sum_{j=1}^{n} f_{ij}x_j + b_i, \quad i = 1, \cdots, n. \tag{6}$$

简记为

$$Y = FX + B,$$

其中, Y, X, B 是相应的坐标列(与§1的(3)式比较, 或者, 更好地, 与第2章§1的(3)式比较).

2. 欧几里得空间的运动 设($\mathbb{E}$, V, ρ)是个欧几里得(点)空间.不必说, 已经假定了$\mathfrak{K} = \mathbb{R}$.

定义1 任意一个变换$f: \mathbb{E} \to \mathbb{E}$, 如果保持距离不变, 即对任意$\dot{p}$, $\dot{q} \in \mathbb{E}$ 都有

$$\rho(f(\dot{p}), f(\dot{q})) = \rho(\dot{p}, \dot{q}), \tag{7}$$

则称f 是一个**运动**(或者**保距映射**).

在运动f 的定义中并没有先假定f 是仿射变换, 但是, 实际上正如下面将要证明的, f 本身必然有:

定理3 变换f : $\mathbb{E} \to \mathbb{E}$ 是个运动, 当且仅当, f 是个仿射变换, 它的线性部分是U上的正交线性算子.

证明 断言的一个方面差不多就是显然的.事实上, 任意一个以正交线性算子$\mathcal{F}$为线性部分的仿射变换f 必然具有性质(7):

$$\begin{aligned}\rho(f(\dot{p}),f(\dot{q}))&=\rho(f(\dot{p}),f(\dot{p}+\mathbf{v}))=\left\|\overrightarrow{f(\dot{p})f(\dot{p}+\mathbf{v})}\right\|\\&=\|\mathcal{F}(\mathbf{v})\|=\|\mathbf{v}\|=\left\|\overrightarrow{\dot{p}(\dot{p}+\mathbf{v})}\right\|=\|\overrightarrow{\dot{p}\dot{q}}\|=\rho(\dot{p},\dot{q}).\end{aligned}$$

特别地, 我们应该注意到, 任何一个平移都是个运动.

关键在于反命题的证明. 我们把这个证明拆成几段.

1) 简单地验证即可表明, 任意两个运动的乘积仍然是个运动. 设f 是个运动, $\dot{o}$ 是个固定点, $\dot{p}=f(\dot{o})$, 而$\mathbf{a}=\overrightarrow{oo'}$ 且$t_{\mathbf{a}}$ 是加上向量$\mathbf{a}$的平移. 那么, $g=t_{\mathbf{a}}^{-1}\cdot f$ 仍然是个运动. 因为

$$g(\dot{o})=t_{\mathbf{a}}^{-1}(f(\dot{o}))=t_{\mathbf{a}}^{-1}(\dot{o}')=\dot{o},$$

所以, $f=t_{\mathbf{a}}g$, 其中$g(\dot{o})=\dot{o}$, 也就是说, 任何一个运动都是一个平移和一个保持$\dot{o}$ 点不动的运动的乘积. 我们只要能证明g是个具有正交线性部分的仿射变换就够了.

2) 进而, 设g是个运动, $g(\dot{o})=\dot{o}$. 我们定义一个变换$\mathcal{G}:V\to V$, 只要令$\mathcal{G}\mathbf{x}=\overrightarrow{\dot{o}g(\dot{o}+\mathbf{x})}$, 也就是

$$g(\dot{o}+\mathbf{x})=\dot{o}+\mathcal{G}\mathbf{x}. \tag{8}$$

变换$\mathcal{G}$ 有性质

$$\mathcal{G}\mathbf{0}=\mathbf{0},\quad \|\mathcal{G}\mathbf{x}-\mathcal{G}\mathbf{y}\|=\|\mathbf{x}-\mathbf{y}\|. \tag{9}$$

实际上, $g(\dot{o})=\dot{o}\Rightarrow\mathcal{G}\mathbf{0}=\mathbf{0}$. 现在设$\dot{p}=\dot{o}+\mathbf{x}$, $\dot{q}=\dot{o}+\mathbf{y}$, 那么, $\rho(\dot{p},\dot{q})=\|\mathbf{y}-\mathbf{x}\|$, 这是因为$\dot{q}=\dot{p}+\mathbf{y}-\mathbf{x}$ 和$\overrightarrow{\dot{p}\dot{q}}=\mathbf{y}-\mathbf{x}$. 又因为$g$是个运动, 所以, $\rho(g(\dot{p}),g(\dot{q}))=\rho(\dot{p},\dot{q})=\|\mathbf{y}-\mathbf{x}\|$. 于是, 由(8)得出, $g(\dot{p})=\dot{o}+\mathcal{G}\mathbf{x},g(\dot{q})=\dot{o}+\mathcal{G}\mathbf{y}$, 从而有$\rho(g(\dot{p}),g(\dot{q}))=\|\mathcal{G}\mathbf{y}-\mathcal{G}\mathbf{x}\|$. 于是, 这就给出了(9)式.

设$\mathbf{y}=\mathbf{0}$, 我们就特别地得到包括了(3)在内的

$$\|\mathcal{G}\mathbf{x}\|=\|\mathbf{x}\|. \tag{10}$$

3) 变换$\mathcal{G}$ 保持纯量乘积不变, 即

$$(\mathcal{G}\mathbf{x}|\mathcal{G}\mathbf{y})=(\mathbf{x}|\mathbf{y}). \tag{11}$$

事实上, 按照(9)

$$\begin{aligned}\|\mathbf{x}\|^2-2(\mathbf{x}|\mathbf{y})+\|\mathbf{y}\|^2&=(\mathbf{x}-\mathbf{y}|\mathbf{x}-\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|^2=\|\mathcal{G}\mathbf{x}-\mathcal{G}\mathbf{y}\|^2\\&=(\mathcal{G}\mathbf{x}-\mathcal{G}\mathbf{y}|\mathcal{G}\mathbf{x}-\mathcal{G}\mathbf{y})=\|\mathbf{x}\|^2-2(\mathcal{G}\mathbf{x}|\mathcal{G}\mathbf{y})+\|\mathcal{G}\mathbf{y}\|^2.\end{aligned}$$

把这个关系式再加上(10)就得到了(11)式.

4) 变换$\mathcal{G}$是线性的, 事实上, 我们设$\mathbf{z}=\mathbf{x}+\mathbf{y}$, 则$\|\mathbf{z}-\mathbf{x}-\mathbf{y}\|^2=0$. 把这个等式改写得更详细些, 就得到

$$\|\mathbf{z}\|^2+\|\mathbf{x}\|^2+\|\mathbf{y}\|^2-2(\mathbf{z}|\mathbf{x})-2(\mathbf{z}|\mathbf{y})+2(\mathbf{x}|\mathbf{y})=0.$$

由这个等式以及(10), (11), 有

$$\|\mathcal{G}\mathbf{z}\|^2+\|\mathcal{G}\mathbf{x}\|^2+\|\mathcal{G}\mathbf{y}\|^2-2(\mathcal{G}\mathbf{z}|\mathcal{G}\mathbf{x})-2(\mathcal{G}\mathbf{z}|\mathbf{y})+2(\mathcal{G}\mathbf{x}|\mathcal{G}\mathbf{y})=0,$$

这等价于等式$\|\mathcal{G}\mathbf{z}-\mathcal{G}\mathbf{x}-\mathcal{G}\mathbf{y}\|=0$, 也就是$\mathcal{G}\mathbf{z}-\mathcal{G}\mathbf{x}-\mathcal{G}\mathbf{y}=\mathbf{0}$. 由此可见

$$\mathcal{G}(\mathbf{x}+\mathbf{y})=\mathcal{G}\mathbf{x}+\mathcal{G}\mathbf{y}.$$

等式$\mathcal{G}(\lambda\mathbf{x})=\lambda\mathcal{G}\mathbf{x}$ 可以类似地加以证明.

5) 证明的结尾. 由(8)和前面3), 4)中已经证明过的, 可以推出, g是个仿射变换, 它的线性部分$\mathcal{G}$ 是个正交的线性算子. □

我们把第1目定理2的证明中的推理精确化, 得

定理4 *设f 是欧几里得空间$(\mathbb{E}, V, \rho)$的一个运动, 它的线性部分$\mathcal{F}$ 是个正交线性算子. 那么, 必有V 的对于$\mathcal{F}$ 不变的正交的子空间的直和分解*

$$V=L\oplus L^{\perp} \tag{12}$$

和一个点$\dot{o}\in\mathbb{E}$, 使得对任意$\mathbf{x}\in L$ 都有$\mathcal{F}\mathbf{x}=\mathbf{x}$, 而$f=t_{\mathbf{a}}\cdot g$, $\mathbf{a}\in L$ 且$g(\dot{o})=\dot{o}$.

证明 用L 代表V 中所有在$\mathcal{F}$ 作用之下保持不变的向量的集合.显然, 它是个对于$\mathcal{F}$ 不变的V 的向量子空间.正如我们已经知道的(第3章§1的定理7), $L^{\perp}$ 同样也是$\mathcal{F}$不变的, 并且有分解式(12).

在$\mathbb{E}$ 中任取一点$\dot{o}'$ 并把f 表示成$f=t_{\mathbf{a}'}\cdot g'$, $g'(\dot{o}')=\dot{o}'$. 当用点$\dot{o}=\dot{o}'+\mathbf{x}$ 替换点$\dot{o}'$ 时, 可以引出$\mathbf{a}=\mathbf{a}'+(\mathcal{F}-\mathcal{E})\mathbf{x}$ (定理2), 而g' 引出g, 且$g(\dot{o})=\dot{o}$.

令$\mathbf{a}'=\mathbf{b}+\mathbf{c}, \mathbf{x}=\mathbf{y}+\mathbf{z}$, 其中$\mathbf{b},\mathbf{y}\in L$, 而$\mathbf{c},\mathbf{z}\in L^{\perp}$. 我们来选择适当形式的向量$\mathbf{x}$. 线性算子$\mathcal{F}-\mathcal{E}$ 限制在$L^{\perp}$ 上, 它的核为零, 因为$L\cap L^{\perp}=0$. 所以$\mathcal{F}-\mathcal{E}|_{L^{\perp}}$ 是非退化的. 特别地, 这意味着, 存在这样一个向量$\mathbf{z}\in L^{\perp}$, 使得$(\mathcal{F}-\mathcal{E})\mathbf{z}=-\mathbf{c}$,从而

$$\mathbf{a}=\mathbf{b}+\mathbf{c}-\mathbf{c}=\mathbf{b}\in L.$$

□

3. 保距变换群 因为所有的仿射变换构成一个群(定理1)且所有的正交线性算子也构成群(第3章§2的定理2), 所以, 按照定理1, 欧几里得空间$\mathbb{E}$ 的所有运动也同样是个群. 我们称它是空间$\mathbb{E}$ 的**保距群**, 并用符号Iso($\mathbb{E}$) 来代表它. 因为同一维数的两个欧几里得空间是同构的(§2定理1). 所以, 精确到同构, 对于每一个维数而言只有一个保距群. 显然, Iso($\mathbb{E}$) 是仿射群Aff($\mathbb{E}$)的一个子群. 在Iso($\mathbb{E}$) 中包含了平移子群T, 后者同构于向量空间V 的加法子群. 在一个固定点$\dot{o}\in\mathbb{E}$ 处保持不变的所有的运动

的子群同构于正交群$O(n)$, $n=\dim\mathbb{E}$. 如果$\{\dot{o};\mathbf{e}_1,\cdots,\mathbf{e}_n\}$是$\mathbb{E}$的一个直角坐标系, 那么, 运动$f$可写成形如

$$Y=FX+A, \tag{13}$$

其中$X=[x_1,\cdots,x_n]$, $Y=[y_1,\cdots,y_n]$分别是点$\dot{p}$和点$f(\dot{p})$的坐标列, $A=[a_1,\cdots,a_n]$是向量$\mathbf{a}\in V$的坐标列, $\mathbf{a}$对应平移$t_{\mathbf{a}}$, F是正交矩阵.

如果, $F\in SO(n)$, 也就是$\det F=1$,那么, 称f是个**固有运动**.空间$\mathbb{E}$的所有的固有运动的群用$\mathrm{Iso}_+(\mathbb{E})$来代表(话又说回来了, 我们不会再用到它).

保距群的元素, 也就是所说的运动, 会经常在几何学和物理学中遇到, 所以, 我们停下来, 解释一下n不太大的情形是很有意义的.

n= 1情形 按照一般公式(13), 有

$$y=\varepsilon x+a, \tag{14}$$

其中$\varepsilon=\pm1$ (1维的正交线性算子), 而a是某个常数, 它对应一个平移. 如果$\varepsilon=1$, 那么, 我们得到平移直线. 如果$\varepsilon=-1$, 那么, (14)的另外一种写法

$$y-a/2=-(x-a/2),$$

引出一个想法去选择新的原点$x=x'+a/2, y=y'+a/2$.现在, 公式$y'=-x'$表明, 我们就有了一条对于某个原点$\dot{o}'$的**反射**(对称)直线.

n= 2 情形 选择直角坐标系$\{\dot{o};\mathbf{e}_1,\mathbf{e}_2\}$, 在它之下, 运动$f$的线性部分$\mathcal{F}$引出一个规范型(第3章, §3的定理10), 可以看出, f的坐标的描述形式可以归结为下列之一

$$\begin{aligned}&1)\ x'=x+a, &&2)\ x'=x+a, &&3)\ x'=x\cos\varphi-y\sin\varphi+a,\\ &\quad y'=y+b; &&\quad y'=-y+b; &&\quad y'=x\sin\varphi+y\cos\varphi+b.\end{aligned}$$

在情形1), 我们得到一个加向量$a\mathbf{e}_1+b\mathbf{e}_2$的平移.在情形2), 需要把坐标原点变成$\dot{o}'=(\dot{o},-b/2)$, 也就是引进新的坐标系$\xi,\eta$:

$$x=\xi\,(x'=\xi'), \qquad y=\eta\,(y'=\eta'+b/2).$$

在此之后, 2)的公式采用

$$\xi'=\xi+a, \quad \eta'=-\eta,$$

形式.

在情形3), 当$\varphi\neq0$时, 变换坐标原点$\dot{o}'=(x_0,y_0)$, 其中x_0,y_0由方程组

$$\begin{aligned}x_0\cos\varphi-y_0\sin\varphi+a&=x_0,\\ x_0\sin\varphi+y_0\cos\varphi+b&=y_0\end{aligned}$$

决定.几何上, 这意味着, $f(\dot{o}') = \dot{o}'$. 点$\dot{o}'$ 的存在性可由定理4给出.在该定理的表达中$L = 0$ (在$\mathcal{F}$ 作用下没有不动点), 且$t_{\mathbf{a}} = e$, 所以$f = g$ 是个纯粹的旋转. 如果进行形式地推演, 那么, 就引入新的坐标ξ, η:

$$x = \xi + x_0 \quad (x' = \xi' + x_0),$$
$$y = \eta + y_0 \quad (y' = \eta' + y_0),$$

这之后, 3)的公式就具有

$$\xi' = \xi \cos\varphi - \eta \sin\varphi,$$
$$\eta' = \xi \sin\varphi + \eta \cos\varphi$$

形式了.

这样一来, 就证明了

定理5 平面上的任何一个固有运动或者是个平移或者是个围绕某个固定点的旋转.可见, 保持某点不动的固有运动就是绕这个点的一个旋转. 平面的非固有运动可归结为一个相对于某条直线(在我们这里就是横坐标轴)的反射以及沿这条直线的一个平面平移.由非固有运动至少存在一个不动点, 即可推出, 有一条直线整个地由不动点组成.

$n=3$ 情形 再次依据第3章§3的定理10, 我们力求在3维欧几里得空间$\mathbb{E}$中选择那样一个直角坐标系$\{\dot{o}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, 使得运动$f$ 的线性部分在这个坐标之下是个规范型, 此时, 对于f 的坐标描述有下列的可能性:

1) $x' = x + a$, $y' = y + b$, $z' = z + c$;

2) $x' = x\cos\varphi - y\sin\varphi + a$, $y' = x\sin\varphi + y\cos\varphi + b$, $z' = z + c$;

3) $x' = x + a$, $y' = y + b$, $z' = -z + c$;

4) $x' = x\cos\varphi - y\sin\varphi + a$, $y' = x\sin\varphi + y\cos\varphi + b$, $z' = -z + c$.

在情形1), 我们得到一个加向量$a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3$ 的平移.

在情形2), 当$\varphi \neq 0$ 时, 类似于在平面上的作法, 我们(经过坐标原点的变换$\dot{o}' = (x_0, y_0, 0)$)得到公式

$$\xi' = \xi \cos\varphi - \eta \sin\varphi,$$
$$\eta' = \xi \sin\varphi + \eta \cos\varphi,$$
$$\mu' = \mu + c.$$

可见, f 是个沿直线$\dot{o}'\mu$ 加上向量$(0, 0, c)$的平移并以这条直线为轴旋转φ 角. 这就得到了一个在力学中所说的**螺旋运动**(把螺母拧到螺栓上去).

在情形3), 采用新的坐标ξ,η,μ :

$$\begin{aligned} x &= \xi, & x' &= \xi', \\ y &= \eta, & y' &= \eta', \\ z &= \mu + c/2; & z' &= \mu' + c/2, \end{aligned}$$

我们导出公式

$$\xi' = \xi + a, \qquad \eta' = \eta + b \qquad \mu' = -\mu,$$

这表明, f 可以被刻画成对于平面$\Pi = \dot{o}\,\xi\,\eta$ 的反射再加上与该平面共面的向量$(a,b,0)$所施的平移.

在情形4), 它是情形2)和3)的组合, 公式可以化成

$$\begin{aligned} \xi' &= \xi\cos\varphi - \eta\sin\varphi, \\ \eta' &= \xi\sin\varphi + \eta\cos\varphi, \\ \mu' &= -\mu, \end{aligned}$$

由此可以得出, f 是个对于$\dot{o}\,\xi\,\eta$ 平面的反射, 再绕$\dot{o}\,\mu$ 轴旋转φ 角.

定理6 *3维欧几里得空间$\mathbb{E}$ 的固有运动f 必然是螺旋运动, 即, f 可以被刻画成沿某条直线平移且绕此直线旋转(螺旋运动包括纯粹的平移和纯粹的旋转).*

非固有运动是一个对于某个平面Π 的反射, 并且, 或者再施以加上一个与同一个Π 共面的向量的平移, 或者再绕一个与平面Π 垂直的直线旋转φ 角(当$\varphi = \pi$ 时得到相对于点的对称).

作为特殊情形, 可以由定理6推出欧拉定理(1776年), 按照这个定理, 所有的具有一个不动点$\dot{o}$ 的刚体位移都可以把自己表示成绕某个轴(通过点$\dot{o}$)的旋转, 还可以推出沙勒定理(1830年), 这是一个关于任意物体的位移都可以通过沿某个方向的无偏位移和一个以轴为方向的一个旋转来实现的定理.

4. 与群对应的线性几何 按着125年前由F. 克莱因(1872年)首先在其"爱尔兰根纲领"中明确叙述过的, 现在已经被大众接受的观点, 几何学可以被理解为给定群之下的不变量的集合. 设Γ是某个集合, 或者是我们还要再讲的点的空间, G是所有$\Gamma \to \Gamma$的双射变换群的某个子群. 符合群G的几何学的目标就是研究Γ中在G的变换作用之下不变的空间图形(或者点的空间构型)的性质.

所有的图形被划分成$G-$叠合图形类, 也就是说, 图形Φ_1 是与图形Φ_2 叠合的(或相等)并记为$\Phi_1 \overset{G}{\sim} \Phi_2$ 是指: 至少有一个元素$g \in G$ 使得$\Phi_2 = g(\Phi_1)$. 由群的公理可以直接推出, 叠合性质是个等价关系, 也就是有下列性质:

i) 自反性(这是因为, $\Phi_1 \overset{G}{\sim} \Phi_1 \Leftrightarrow \Phi_1 = e\,(\Phi_1)\,, e$是群$G$的单位元);

ii) 对称性(这是因为, $\Phi_2 = g(\Phi_1) \Rightarrow \Phi_1 = g^{-1}(\Phi_2) g^{-1} \in G$, 也就是, $\Phi_1 \overset{G}{\sim} \Phi_2 \Leftrightarrow \Phi_2 \overset{G}{\sim} \Phi_1$);

iii) 传递性(这是因为, $\Phi_1 = g(\Phi_2), \Phi_2 = h(\Phi_3) \Rightarrow \Phi_1 = (gh)(\Phi_3)$).

这样一来, 叠合图形的类与类之间是不相交的.在线性几何学中要研究的空间Γ或者是线性(向量)空间或者是由直线派生出来的空间. 最接近的例子是**欧几里得几何学**($\Gamma = \mathbb{E}, G = \mathrm{Iso}(\mathbb{E})$) 和**仿射几何学**($\Gamma = \mathbb{A}, G = \mathrm{Aff}(\mathbb{A})$) . 这两种几何学已经由被研究的图形的集合区分开了. 初等欧几里得几何学, 在平面上处理直线, 角, 三角形, 圆, 等等, 还有不同的图形元素中直线和角的关系. 而仿射几何学以点与点之间的距离为发展前提. 我们停下来研究一下在仿射几何和欧几里得几何中图形的最简单的性质.

定理7 *设$(\mathbb{E}, V, \rho)$ 是一个欧几里得空间. 任意两个平面$\Pi, \Pi' \in \mathbb{E}$ 是叠合的($G = \mathrm{Iso}(\mathbb{E})$) , 当且仅当, $\dim\ \Pi = \dim\ \Pi'$. 特别地, 所有的点都是叠合的. 在仿射空间$(\mathbb{A}, V)$ 和群$G = \mathrm{Aff}(\mathbb{A})$ 的情形, 有同样命题成立.*

证明 事实上, 如果$\Pi = \dot{p} + U, \Pi' = \dot{p}' + U'$且$f(\Pi) = \Pi', f \in G$, 那么, $Df(U) = U'$, 又因为$\det\ f := \det\ Df \neq 0$, 所以$\dim\ U = \dim\ U'$, 从而, 按定义$\dim\ \Pi = \dim\ \Pi'$.

反过来, 设$\dim\ \Pi = \dim\ \Pi' = m$. 在U中选取一个标准正交基底$(\mathbf{e}_1, \cdots, \mathbf{e}_m)$ (相应地, 在U' 中选取一个标准正交基底$(\mathbf{e}'_1, \cdots, \mathbf{e}'_m)$), 然后把它扩展成整个空间$V$的标准正交基底$(\mathbf{e}_i)$ (相应地, $(\mathbf{e}'_i)$). 于是, 必存在正交线性变换$\mathcal{F}: V \to V$ 使得$\mathcal{F}\mathbf{e}_i = \mathbf{e}'_i; i = 1, \cdots, n = \dim V$. 满足$f(\dot{p}) = \dot{p}'$ 且$Df = \mathcal{F}$ 的运动f 就把Π 变成Π' .

在仿射空间$(\mathbb{A}, V)$ 的情形, 推理可以类似地完成. 并不需要为标准正交性和正交性操心. □

注意到在定理7的阐述中, 点的叠合性还可以用另外的语言来表达: 群G的作用在空间$\mathbb{E}$ (相应地$\mathbb{A}$)的点上是可以传递的. 传递性是G和与其相符合的几何学的一个最重要的性质. 如果没有这一点, 我们就丧失了不同图形之间"比较"的可能性. 在仿射几何情形, 群Aff $(\mathbb{A})$ 具有更加有力得多的性质.

定理8 *在仿射几何中, $(\mathbb{A}, V)$ 的任意一个处于普通位置的$m+1$ 点的点组$\{\dot{p}_0, \cdots, \dot{p}_m\}$ 和任意一个也处于普通位置的点组$\{\dot{p}'_0, \cdots, \dot{p}'_m\}, 0 \leqslant m \leqslant n$, 是可以叠合的.*

证明 把给定的点组扩充成仍然处于普通位置的点组$\{\dot{p}_0, \cdots, \dot{p}_n\}$和$\{\dot{p}'_0, \cdots, \dot{p}'_n\}$, $n = \dim\ \mathbb{A}$. 按定义, 这意味着, 向量组($\mathbf{e}_1 = \overrightarrow{p_0p_1}, \cdots, \mathbf{e}_n = \overrightarrow{p_0p_n}$) 构成空间$V$ 的一个基底, 而($\mathbf{e}'_1 = \overrightarrow{p'_0p'_1}, \cdots, \mathbf{e}'_n = \overrightarrow{p'_0p'_n}$) 是另一个基底. 于是, 可以找到非退化的线性算子$\mathcal{F}: V \to V$ 使得$\mathcal{F}\mathbf{e}_i = \mathbf{e}'_i$. 令$f(\dot{p}_o + \mathbf{x}) = \dot{p}_0 + \mathcal{F}\mathbf{x}$, 我们就得到了所求的仿射变换$f: \mathbb{A} \to \mathbb{A}$, 它把$\dot{p}_i$ 变成$\dot{p}'_i, i = 0, 1, \cdots, n$. □

明显地, 定理8对于欧几里得几何, 甚至于在m=1的情形, 也不再对了, 因为点对$\dot{p}$, $\dot{q}$ 和点对$\dot{p}', \dot{q}'$的$\mathrm{Iso}(\mathbb{E})$ -叠合必须满足条件$\rho(\dot{p}, \dot{q}) = \rho(\dot{p}', \dot{q}')$. 但是, 这个条件却是充分的, 这可以由定理7的证明看得出来.

仿射自同构映射的几何意义同样可以由下面的推断看出来. 我们考察任意一个双射变换$f: \mathbb{A} \to \mathbb{A}$, 有

$$\overrightarrow{rs} = \lambda\overrightarrow{pq} \Rightarrow \overrightarrow{f(r)f(s)} = \lambda\overrightarrow{f(p)f(q)}, \forall\lambda \in \mathbb{R} \tag{15}$$

(这里的实数域$\mathbb{R}$ 可以用任意其他域来代替). 几何上, 这意味着, f 把共线的点变成仍然共线的点, 即, 把仿射直线变成另外一条仿射直线.

设$\mathcal{F}(\overrightarrow{pq}) = \overrightarrow{f(p)f(q)}$, 我们看到(当$\lambda = 1$ 时), 变换$\mathcal{F}: V \to V$ 的结果与点$\dot{r}, \dot{s} \in \mathbb{A}$ 的选择无关, 只要是$\overrightarrow{rs} = \overrightarrow{pq}$, 就由向量$\overrightarrow{pq}$本身决定了. 可以证明, 定义的这个映射$\mathcal{F}$ 是线性的. 条件$\mathcal{F}(\lambda\mathbf{v}) = \lambda\mathcal{F}(\mathbf{v})$ 可以由定义和条件(15)推导出来. 任意两个向量$\mathbf{u}, \mathbf{v} \in V$ 可以表示成$\mathbf{u} = \overrightarrow{pq}, \mathbf{v} = \overrightarrow{qr}$, 其中$\dot{p}$, $\dot{q}$, $\dot{r} \in \mathbb{A}$. 所以, $\mathbf{u}+\mathbf{v} = \overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$, 而且

$$\mathcal{F}(\mathbf{u}+\mathbf{v}) = \overrightarrow{f(p)f(r)} = \overrightarrow{f(p)f(q)} + \overrightarrow{f(q)f(r)} = \mathcal{F}\mathbf{u} + \mathcal{F}\mathbf{v}.$$

可见, $\mathcal{F}$ 是V 上的线性算子. 对任意$\dot{q} = \dot{p} + \mathbf{x}$, 我们有

$$f(\dot{q}) = f(\dot{p}) + \overrightarrow{f(p)f(q)}, \quad \overrightarrow{f(p)f(q)} = \mathcal{F}(\overrightarrow{pq}) = \mathcal{F}\mathbf{x},$$

从而$f(\dot{p}+\mathbf{x}) = f(\dot{p}) + \mathcal{F}\mathbf{x}$, 也就是说, f 具有性质(15), 显然是个仿射变换.

反之亦然: 如果f 是个以$\mathcal{F}$ 为线性部分的仿射自同构, 而且只要$\overrightarrow{rs} = \lambda\overrightarrow{pq}$, 则必有$\mathcal{F}(\overrightarrow{rs}) = \lambda\mathcal{F}(\overrightarrow{pq})$. 而$f(\dot{s}) = f(\dot{r} + \overrightarrow{rs}) = f(\dot{r}) + \mathcal{F}(\overrightarrow{rs})$, 因此, $\mathcal{F}(\overrightarrow{rs}) = \overrightarrow{f(r)f(s)}$. 我们就证明了

定理9 双射映射$f: \mathbb{A} \to \mathbb{A}$ 的性质(15)是仿射映射的固有特征.

现在看一个特殊情形, $\dot{p}, \dot{q}, \dot{r}$ 三点在同一条直线上(通常说是**共线**三点), 且$\dot{p} \neq \dot{q}$. 于是, 可以找到一个数λ 使得

$$\overrightarrow{pr} = \lambda\overrightarrow{pq}. \tag{15'}$$

定义2 公式(15′)中的数λ 被称为共线三点$\dot{p}, \dot{q}, \dot{r}$ 的**最简比例**, 并记为$[\dot{p}, \dot{q}, \dot{r}]$.

显然, 由定理9可以得到

推论 空间$\mathbb{A}$ 的仿射变换f 保持点的共线性质并保持共线三点组的最简比例.

公式(15′)还可以写成"加法型"公式$\dot{r} - \dot{p} = \lambda(\dot{q} - \dot{p})$, 这直接意味着, 在直线$\prod_{\dot{p},\dot{q}}$上点$\dot{r}$ 可以写成$\dot{r} = (1-\lambda)\dot{p} + \lambda\dot{q}$ 的样子. 特别地, 在仿射几何中, 处于线段$\dot{p}\dot{q}$ 内部的点$\dot{r}, 0 < \lambda < 1$, 内点的比例是有意义的. 前面我们已经指出过, 线段的长度是欧几里得几何的概念, 而线段的中点是个仿射概念.

5. 欧几里得空间的仿射变换 在我们的周围世界到处呈现出仿射变换作用的影响. 最简单的例子是松紧带的压缩. 我们还要更仔细的考察某些事实. 在第4章§2, 我们已经约定, 把量$v_n = |\det(a_{ij})|_1^n$ 理解为以$\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_n$ 为边的平行六面体 $P(\dot{o}\dot{p}_1, \cdots, \dot{o}\dot{p}_n)$ 的体积v_n , 其中(a_{ij}) 是由欧几里得空间V 中的标准正交基底$(\mathbf{f}_1, \cdots, \mathbf{f}_n)$ 向基底$\mathbf{e}_i = \overrightarrow{op_i}, i = 1, \cdots, n$ 的转换矩阵. 另一方面, 如果g是以$\mathcal{G}$ 为线性部分的仿射自同构, 那么, 向量$\mathcal{G}\,\mathbf{e}_1, \cdots, \mathcal{G}\mathbf{e}_n$ (严格地说, 是与这些向量匹配的线段)构成的平行六面体的体积是$v_n' = |\det\,(b_{jk})|$, 其中矩阵(b_{jk})可以按下面的规则计算出来. 设

$$\mathcal{G}\mathbf{f}_i = \sum_{j=1}^{n} g_{ji}\mathbf{f}_j.$$

那么,

$$\sum_j b_{jk} f_j := \mathcal{G}\mathbf{e}_k = \sum_i a_{ik}\mathcal{G}\mathbf{f}_i = \sum_i a_{ik}\sum_j g_{ji}f_j = \sum_j(\sum_i g_{ji}a_{ik})f_j,$$

也就是$b_{jk} = \sum_i g_{ji}a_{ik}$, 从而

$$B = GA.$$

进而有

$$v_n' = |\det(b_{jk})| = |\det G|\cdot v_n = |\det\ g|\cdot v_n.$$

我们就引出了下列结论

定理10　在n 维欧几里得空间的仿射变换之下, n 个向量构成的平行六面体的体积就被乘以一个转换行列式的绝对值. 换言之, 在仿射变换之下, 平行六面体的体积比保持不动.

同样地, 也可以归纳出对于欧几里得空间其他任意形体的体积.

下列命题有直观几何意义.

定理11　n 维欧几里得空间$(\mathbb{E}, V)$中所有非退化的仿射变换f 都是个乘积

i) 施以某个向量的平移;

ii) 保持某个固定点$\dot{o}$ 不动的旋转;

iii) 沿着相交于点$\dot{o}$ 的相互垂直的轴的n 次压缩(或伸张)的复合仿射变换h.

证明　实际上, 按照定理2, $f = t_{\mathbf{a}}\cdot g$, 其中$g(\dot{o}) = \dot{o}$, 对某个点$\dot{o}$. 如果$\mathcal{G}$ 是变换g的线性部分, 那么, 按第3章§3的定理15, $\mathcal{G} = \mathcal{D}\mathcal{H}$, 其中, $\mathcal{D}$ 是V 上的正交线性算子, 而$\mathcal{H}$ 是个正定的对称矩阵. 相应于第3章§3的定理6, 我们在$\mathbb{E}$ 中选择直角坐标系$\{\dot{o};\mathbf{e}_1,\cdots,\mathbf{e}_n\}$, 此基底之下, 算子$\mathcal{H}$ 有规范形式

$$\mathcal{H}e_i = \lambda_i\mathbf{e}_i,\quad \lambda_i > 0,\quad i = 1,\cdots,n.$$

于是有

$$f = t_{\mathbf{a}}\cdot d\cdot h,\quad d(\dot{o}+\mathbf{x}) = \dot{o}+\mathcal{D}\mathbf{x},\quad h(\dot{o}+\mathbf{x}) = \dot{o}+\mathcal{H}\mathbf{x}, \tag{16}$$

因此, d是空间$\mathbb{E}$ 的一个旋转, 而h是仿射变换, 它可以写成一个乘积形式

$$h = h_1h_2\cdots h_n. \tag{17}$$

这里, h_k 是以$\mathcal{H}_k$ 为线性部分的仿射变换:

$$\mathcal{H}_k\mathbf{e}_i = \mathbf{e}_i,\quad \text{当}i\neq k\text{时};\quad \mathcal{H}_k\mathbf{e}_k = \lambda_k\mathbf{e}_k.$$

公式(16)和(17)给出了所要求的仿射变换f 的分解. □

6. 凸集　回顾一下点的重心组合的定义(§1的第5目)

$$\dot{p} = \lambda_0\dot{p}_0 + \lambda_1\dot{p}_1 + \cdots + \lambda_m\dot{p}_m,\quad \lambda_0+\lambda_1+\cdots+\lambda_m = 1, \tag{18}$$

和重心坐标的定义, 我们现在可以看出, 当m=1时, 点$\dot{p} = \lambda_0\dot{p}_0 + \lambda_1\dot{p}_1$, $\lambda_0 + \lambda_1 = 1$, 就取遍了整个直线$\Pi_{\dot{p}_0,\dot{p}_1}$. 当$m$=2时, 点$\dot{p} = \lambda_0\dot{p}_0 + \lambda_1\dot{p}_1 + \lambda_2\dot{p}_2$, $\lambda_0 + \lambda_1 + \lambda_2 = 1$, $\lambda_i > 0$ 就取遍了以$\dot{p}_0, \dot{p}_1, \dot{p}_2$ 为顶点的开三角形.

事实上, 三角形的内点$\dot{p}$ 就是线段$\dot{p}_0\dot{q}$ 的一个内点, 而$\dot{q}$是线段$\dot{p}_1\dot{p}_2$ 的一个内点. 即, 有

$$\dot{p} = \lambda_0\dot{p}_0 + \lambda\dot{q}\,, \quad \lambda_0 + \lambda = 1, \quad \lambda_0 > 0\,,\ \lambda > 0,$$

$$\dot{q} = \alpha_1\dot{p}_1 + \alpha_2\dot{p}_2\,, \quad \alpha_1 + \alpha_2 = 1, \quad \alpha_1 > 0\,,\ \alpha_2 > 0.$$

可见,

$$\dot{p} = \lambda_0\dot{p}_0 + \lambda(\alpha_1\dot{p}_1 + \alpha_2\dot{p}_2) = \lambda_0\dot{p}_0 + \lambda_1\dot{p}_1 + \lambda_2\dot{p}_2,$$

其中$\lambda_1 = \lambda\alpha_1 > 0, \lambda_2 = \lambda\alpha_2 > 0$, 且$\lambda_0 + \lambda_1 + \lambda_2 = 1$ (直接计算或由§1习题4).

反之, 如果$\dot{p} = \lambda_0\dot{p}_0 + \lambda_1\dot{p}_1 + \lambda_2\dot{p}_2$及$\lambda_0 + \lambda_1 + \lambda_2 = 1$, $\lambda_i > 0$, $i = 0$, 1 , 2 , 所以, $\dot{p} = \lambda_0\dot{p}_0 + \lambda\dot{q}$, 其中$\lambda = \lambda_1 + \lambda_2 > 0$, $\lambda_0 + \lambda = 1$, 进而$\dot{q} = \alpha_1\dot{p}_1 + \alpha_2\dot{p}_2, \alpha_1 = \lambda_1/(\lambda_1 + \lambda_2), \alpha_2 = \lambda_2/(\lambda_1 + \lambda_2)$, 因此, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_1 + \alpha_2 = 1$. 也就是说, $\dot{q}$ 是线段$\dot{p}_1\dot{p}_2$ 的内点, 而$\dot{p}$ 是以$\dot{p}_0$, $\dot{p}_1, \dot{p}_2$ 为顶点的三角形的一个内点(图8).

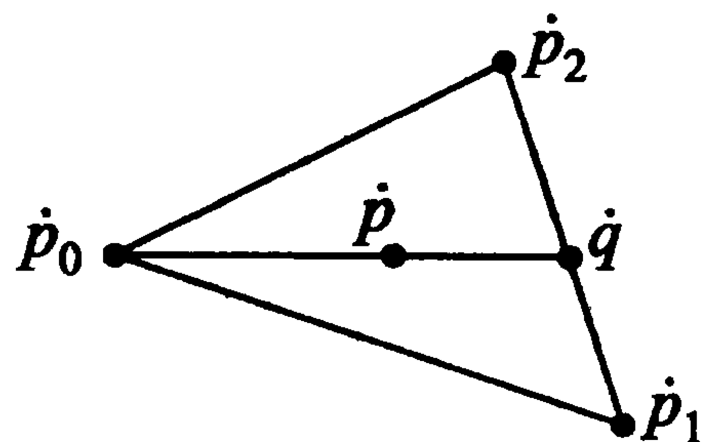


图8

用类似的方式推断m=3的情形, 我们就得到了四面体, 而对任意$m \leqslant n$ 的情形, 就得到单纯形. 于是就有了下面的

定义3 称所有形如(18)的具有正的重心坐标λ_0, λ_1 , $\cdots$, λ_m 的点的集合是以$\dot{p}_0$, $\dot{p}_1$, $\cdots$,$\dot{p}_m$ 为顶点的m维**开单纯形**. 对于点系$\dot{p}_0$, $\dot{p}_1$, $\cdots$,$\dot{p}_m$ 的非负的重心坐标, 相应地, 称为是以$\dot{p}_0$, $\dot{p}_1$, $\cdots$,$\dot{p}_m$ 为顶点的**闭单纯形**.

定理12 *在仿射自同构之下, 任何单纯形的像仍然是个单纯形. 所有的m维的单纯形在仿射几何里都是叠合的.*

证明 这差不多就是显然的. 设$f : \mathbb{A} \to \mathbb{A}$ 是个仿射变换, 它的线性部分是线性算子$\mathcal{F} : V \to V$. 把f 作用到$\lambda_i \geqslant 0$ 的等式(18)的两部分上, 并利用§1的命题2中的i), 我们就导出等式

$$f(\dot{p}) = \lambda_0 f(\dot{p}_0) + \lambda_1 f(\dot{p}_1) + \cdots + \lambda_m f(\dot{p}_m),$$

这意味着, $f(\dot{p})$ 是以$f(\dot{p}_0)$, $f(\dot{p}_1)$, $\cdots$, $f(\dot{p}_m)$ 为顶点的单纯形的一个点. 定理的最后面的论断只是定理8的论断稍加改动而已. □

定义4 设$(\mathbb{A}\,,\ V)$ 是个仿射空间, 说子集$M \subset \mathbb{A}$ 是**凸的**, 如果它能整个地包含它的任意两个点$\dot{p}$, $\dot{q}$ 所形成的线段$\dot{p}\dot{q}$.

单纯形是凸集的一个重要例子. 明显地, 任意多个凸集的交集仍然是个凸集.

定义5 包含一个给定的集合M的所有凸集的交集被称为集合M的**凸闭包**且用$C(M)$代表.

显然, $C(M) = M$ 的充要条件是M本身是个凸集. 严格地说, 以$\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_m$ 为顶点的m 维单纯形就是所给的这组点的凸闭包.

命题1 设M是个凸集, $\dot{p} \in \mathbb{A}$. 那么

$$C(M \cup \dot{p}) = \cup \dot{p}\dot{q} \ , \quad q \in M.$$

证明 按定义, 线段$\dot{p}\dot{q}$, $\dot{q} \in M$ 属于每一个包含M和$\dot{p}$ 的凸集, 可见 $\bigcup\limits_{\dot{q} \in M} \dot{p}\dot{q} \subset C(M \cup \dot{p})$.

反之, 结论可以由$\cup \dot{p}\dot{q}$, $\dot{q} \in M$ 的凸性推出来, 我们就来检验这一点.

设$\dot{q}_1$, $\dot{q}_2 \in M$, 那么任意点$\dot{r}_1 \in \dot{p}\dot{q}_1$, $\dot{r}_2 \in \dot{p}\dot{q}_2$ 对应$\dot{r} \in \dot{r}_1\dot{r}_2$, 我们可以指出, $\dot{r} \in \dot{p}\dot{q}$, 其中$\dot{q}$ 是M中的某个点. 首先, 设点$\dot{p}$, $\dot{q}_1$, $\dot{q}_2$ 不在一条直线上, 那么, 它们属于自己的仿射闭包, 2维平面$\Pi = A(\dot{p}$, $\dot{q}_1$, $q_2)$, 我们有理由在它上面运用普通的初等几何学. 特别地, 我们注意到, 直线$\Pi_{\dot{p},\dot{r}}$ 与线段$\dot{q}_1$ $\dot{q}_2$ 相交于某一点$\dot{q}$. 由于M 的凸性, 可知$\dot{q} \in M$ (图9). 在这种情形$\dot{r} \in \dot{p}\dot{q}$, 这就完全证明了断言. 如果$\dot{p}$, $\dot{q}_1$, $\dot{q}_2$ 位于同一条直线上, 那么, 可以简单地由$\dot{q}_1$, $\dot{q}_2$ 取一点作为q. □

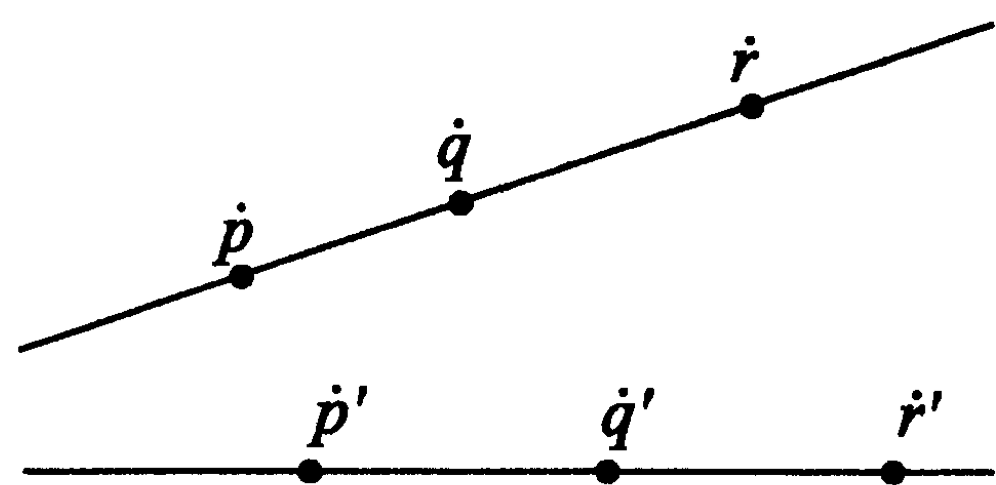


图9

定理13 *在有限个点的点系$\dot{p}_i$, $i = 0$, 1 , $\cdots$, m 的凸闭包$S = C(\dot{p}_0$, $\dot{p}_1$, $\cdots$, $\dot{p}_m)$ 上(也就是在单纯形上)的仿射线性函数f 必然在它的某一个顶点上达到极大值*

$$\max_{\dot{p} \in S} f(\dot{p}) = \max_i f(\dot{p}_i).$$

证明 当m=0时, 定理的断言是平凡的. 然后, 我们对$m > 0$ 用归纳法. 按归纳法假定, 可以认为, 在凸集$M = C(\dot{p}_0$, $\cdots$, $\dot{p}_{m-1})$ 上函数的极大值$\max\limits_{i<m} f(\dot{p}_i)$. 由命题1, 所有的点$\dot{s} \in S$ 必然属于某个线段$\dot{p}_m\dot{q}$, $\dot{q} \in M$, 而且, 这意味着

$$\dot{s} = \dot{p}_m + \lambda\overrightarrow{p_m q}, \qquad 0 \leqslant \lambda \leqslant 1.$$

如果$\mathcal{F}$ 是函数f 的线性部分, 那么,

$$f(\dot{s}) = f(\dot{p}_m) + \lambda\mathcal{F}(\overrightarrow{p_m q}), \mathcal{F}(\overrightarrow{p_m q}) = f(\dot{q}) - f(\dot{p}_m),$$

所以,

$$f(\dot{s}) = (1-\lambda) f(\dot{p}_m) + \lambda f(\dot{q}) \leqslant \max\{f(\dot{p}_m), f(\dot{q})\} \leqslant \max_{i \leqslant m} f(p_i).$$ □

并不复杂的定理13与线性规划分支有密切的关系, 而线性规划有实践意义.

习　　题

1. p元域(p是个素数)上仿射直线的自同构群$A_1(\mathbb{F}_p)$ 的阶数为$p(p-1)$. 问: $A_1(\mathbb{F}_3)$ 同构于怎样的群.

2. 如果欧几里得平面上的一个固有运动f 满足

$$Df = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \qquad f(\dot{o}) = (1,1),$$

试给出f 的几何描述.

3. 试决定4维的欧几里得点空间中固有运动的分类.

§4 带有指数有限度量的空间

1. 指数有限度量 我们已经约定, 带有纯量乘积的空间是指把向量空间V 连同一个固定的二次型

$$q(\mathbf{x}) = f(\mathbf{x},\mathbf{x}) = \sum_{i,j} a_{ij}x_i x_j$$

一起加以研究. 对应于正定型q (通常的或埃尔米特的)的欧几里得空间和埃尔米特(酉)空间已经得到了我们足够详细的研究. 带有由一个不定型q 对应的所谓的指数有限的度量的空间同样也起重要作用. 正如在第1章§4中已经知道的, 通过选择空间V 的适当的基底($\mathbf{e}_i$), 非退化二次型q可取标准形式

$$q(\mathbf{x}) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_n^2 \tag{1}$$

(现在认为基础域是$\mathbb{R}$). 在V 上的纯量乘积就是

$$(\mathbf{x}|\mathbf{y}) := x_1y_1 + \cdots + x_sy_s - x_{s+1}y_{s+1} - \cdots - x_ny_n.$$

为了停留在处理实数领域, 我们只讨论向量$\mathbf{x}$的模(长度)的平方$\|\mathbf{x}\|^2 = (\mathbf{x}|\mathbf{x})$,在$1 \leqslant s \leqslant n-1$ 的时候, 它可能是正的, 也可能取负值. 如果$\|\mathbf{x}\|^2 = 0$, 则说向量$\mathbf{x}$是**迷向的**.

在与向量空间V 相伴的仿射空间$\mathbb{E}$ 中, 点$\dot{p}(x_1,\cdots,x_n)$ 与点$\dot{q}(y_1,\cdots,y_n) \in \mathbb{E}$ 之间的“距离”的平方定义为

$$\rho^2(\dot{p},\dot{q}) = \sum_{i=1}^{s}(y_i - x_i)^2 - \sum_{i=s+1}^{n}(y_i - x_i)^2$$

仍然把二次型$(\mathbf{x}|\mathbf{x})$ 称为向量空间V 上的**度量型**, 而把$\rho^2(\dot{p},\dot{q})$ 称为仿射空间$\mathbb{E}$的一个**度量型**. 当$1 \leqslant s \leqslant n-1$ 时, 说空间$\mathbb{E}$ 是个**伪欧几里得空间**. 当$s=1$ 时, 仍然说它是个**闵可夫斯基空间**(有时把它当作$s=n-1$的情形对待, 这是非常本质的: 用$-q$代替q 即可). 在$n=4$的情形, 闵可夫斯基空间对应物理学上的时空连续统, 后者在与狭义相对论有关所有问题中都起重要作用.

2. 伪欧几里得运动 依照在§3第4目中讲述过的一般概念, 伪欧几里得几何由伪欧几里得运动群决定, 这个群派生出一个平行移动(位移)子群T和某个固定点$\dot{o} \in \mathbb{E}$的稳定子群$O(s, n-s)$ (保持$\dot{o}$ 点不动的子群). 当$s = n$ 时, 有正交群$O(n) = O(n, 0)$. 通常情况下, "伪正交"群$O(s, n-s)$ 由保持二次型(1)的线性算子$\mathcal{F} : V \to V$ 组成. 同样可以说, $O(s, n-s)$ 是型q 的**自同构群**.

选取空间V 的一个规范基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$, 型(1)的矩阵是

$$I_s = \begin{pmatrix} E_s & 0 \\ 0 & -E_{n-s} \end{pmatrix},$$

而算子$\mathcal{F} \in O(s, n-s)$ 的矩阵F满足

$${}^tF \cdot I_s \cdot F = I_s.$$

为了理解这一点, 只需回顾一下二次型在一个给定的基底向另一个基底转换时的矩阵, 在这里是向基底$(\mathcal{F}\mathbf{e}_1, \cdots, \mathcal{F}\mathbf{e}_n)$ 转换. 显然, 就如同在正交群的情形一样, $\det\ \mathcal{F} = \det\ F = \pm 1$. 如果$\det\ \mathcal{F} = 1$, 就说$\mathcal{F}$ 是二次型q 的**固有自同构**, 而伴有$Df = \mathcal{F}$ 的仿射变换$f : \mathbb{E} \to \mathbb{E}$ 则说成是固有的**伪欧几里得运动**. 我们还要注意, 二次型q 的自同构$\mathcal{F}$ 把迷向的向量变成迷向向量. 这是因为$q(\mathcal{F}\mathbf{x}, \mathcal{F}\mathbf{x}) = q(\mathbf{x}, \mathbf{x}) = 0$.

3. 洛伦茨群 正如已经指出过的, 带有符号差为$(1, 3)$ 的非退化的对称矩阵的4维的实向量空间占有特殊地位.

定义1 *群$O(1,3)$称之为***洛伦茨群***并用$\mathbf{L}$代表它.*

在这种情况下, 表达方式

$$\begin{aligned} &V = \langle \mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle, \\ &\mathbf{x} = t\mathbf{e}_0 + x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3, \\ &\|\mathbf{x}\|^2 = q(\mathbf{x}) = t^2 - x_1^2 - x_2^2 - x_3^2 \end{aligned}$$

是标准的.

一种特别的情形, 2维空间中保持度量

$$(\mathbf{u}|\mathbf{u}) = t^2 - x^2$$

的自同构作成的1维的洛伦茨群$\mathbf{L}_1$ 是足够吸引人的. 群$\mathbf{L}_1$ 刻画沿一条直线的物理运动(现在, 在我们这里, x 不是一个向量, 而是向量$\mathbf{u} = t\mathbf{e}_0 + x\mathbf{e}_1$ 的坐标). 显然, 所有的迷向向量都与$\mathbf{e}_0 + \mathbf{e}_1$和$\mathbf{e}_0 - \mathbf{e}_1$ 成比例. 所以, 对于算子$\mathcal{F}$ 来说, 由于它的非退化性, 它只有两种可能

$$\begin{aligned} &\mathcal{F}(\mathbf{e}_0 + \mathbf{e}_1) = \alpha(\mathbf{e}_0 + \mathbf{e}_1), \quad \mathcal{F}(\mathbf{e}_0 - \mathbf{e}_1) = \beta(\mathbf{e}_0 - \mathbf{e}_1), \\ &\mathcal{F}(\mathbf{e}_0 + \mathbf{e}_1) = \alpha(\mathbf{e}_0 - \mathbf{e}_1), \quad \mathcal{F}(\mathbf{e}_0 - \mathbf{e}_1) = \beta(\mathbf{e}_0 + \mathbf{e}_1). \end{aligned}$$

研究这两种可能性中的一个, 比如, 第一个. 有

$$\mathcal{F}\mathbf{e}_0 = \frac{\alpha+\beta}{2}\mathbf{e}_0 + \frac{\alpha-\beta}{2}\mathbf{e}_1,$$
$$\mathcal{F}\mathbf{e}_1 = \frac{\alpha-\beta}{2}\mathbf{e}_0 + \frac{\alpha+\beta}{2}\mathbf{e}_1.$$

算子$\mathcal{F}$ 的矩阵

$$F = \begin{pmatrix} \dfrac{\alpha+\beta}{2} & \dfrac{\alpha-\beta}{2} \\ \dfrac{\alpha-\beta}{2} & \dfrac{\alpha+\beta}{2} \end{pmatrix}$$

的行列式$\det\ F = \alpha\beta$. 我们限于固有的洛伦茨变换的情形, 也就是有$\alpha\beta = 1$. 对于坐标变换, 我们得到

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} \dfrac{\alpha^{-1}+\alpha}{2} & \dfrac{\alpha^{-1}-\alpha}{2} \\ \dfrac{\alpha^{-1}-\alpha}{2} & \dfrac{\alpha^{-1}+\alpha}{2} \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix},$$

从而

$$t' = \frac{\alpha^{-1}+\alpha}{2}\left(t + \frac{\alpha^{-1}-\alpha}{\alpha^{-1}+\alpha}x\right),$$
$$x' = \frac{\alpha^{-1}+\alpha}{2}\left(\frac{\alpha^{-1}-\alpha}{\alpha^{-1}+\alpha}t + x\right).$$

我们引进表达式

$$\frac{\alpha-\alpha^{-1}}{\alpha+\alpha^{-1}} = v = \frac{\alpha^2-1}{\alpha^2+1}. \tag{2}$$

还有一个不至于引起混淆的事实: 在我们这里, 正如前面已经说过的, v 是个纯量, 而不是向量. 所研究的这个量对应物理上的速度, 而速度都直接用符号v 代表. 注意, 我们总有$|v| < 1$, 从而, 由关系式(2)导出来的表达式

$$\alpha^2 = \frac{1-v}{1+v}, \qquad \alpha = \sqrt{\frac{1-v}{1+v}},$$
$$\frac{\alpha+\alpha^{-1}}{2} = \frac{1}{\sqrt{1-v^2}}$$

是有意义的.

最后, 我们得到

$$t' = \frac{t-vx}{\sqrt{1-v^2}}, \qquad x' = \frac{x-vt}{\sqrt{1-v^2}}. \tag{3}$$

这个雅致的公式以把光速取成单位1时的尺度列出来的. 在通常的尺度下, 变换的形式是

$$t' = \frac{t-vx/c^2}{\sqrt{1-v^2/c^2}}, \qquad x' = \frac{x-vt}{\sqrt{1-v^2/c^2}}. \tag{3'}$$

它们对应的二次型是$c^2t^2-x^2$. 为简便计, 我们将使用公式(3). 值得注意的是, 麦克斯韦电学方程并没有改变应用洛伦茨和爱因斯坦变换的结果. 数学家庞加莱首先指出了它的意义, 然后假定, 所有物理定律在洛伦茨变换之下都不应改变(当n=4时). 这应该算是狭义相对论的开始. 我们不停留在物理解释上, 而是讨论公式(3)的进一步的结果. 只要注意到, 在速度v 接近于零(与光速比较, 很小)的情况下, 变换(3)就可以采用伽里略变换

$$t'=t,\qquad x'=x-vt$$

的形式. 但是, 在一般情形, 点的位置可由两个坐标(t, x)——时间和空间来刻画. 我们假定$(x_1,t_1),(x_2,t_1)$ 具有相同的$t=t_1$. 在第一个(不动的)坐标系中对应第二个坐标系中不同的t_1',t_2'. 由此可以得到, 比如说, 长度变换定律

$$x_1'-x_2'=\frac{x_1-vt_1}{\sqrt{1-v^2}}-\frac{x_2-vt_1}{\sqrt{1-v^2}}=\frac{x_1-x_2}{\sqrt{1-v^2}}.$$

反过来, 当$x_1=x_2,t_1\neq t_2$ 时, 我们就得到时间变换定律.

如果f_v 是个由参量v 决定的洛伦茨变换(3), 那么,

$$g_{v_1}\cdot g_{v_2}=g_v.$$

我们来计算参变量(速度)v. 设

$$t'=\frac{t-v_1x}{\sqrt{1-v_1^2}},\qquad x'=\frac{x-v_1t}{\sqrt{1-v_1^2}},$$
$$t''=\frac{t'-v_2x'}{1-v_2^2},\qquad x''=\frac{x'-v_2t'}{\sqrt{1-v_2^2}},$$

就得到

$$t''=\frac{t-v_1x-v_2(x-v_1t)}{\sqrt{1-v_1^2}\cdot\sqrt{1-v_2^2}}=\frac{t-(v_1+v_2)x/(1+v_1v_2)}{\sqrt{1-((v_1+v_2)/(1+v_1v_2))^2}}.$$

这就意味着,

$$v=\frac{v_1+v_2}{1+v_1v_2},$$

这就是最简单的速度复合定律.

4. 真洛伦茨群 所说的由二次型

$$q(\mathbf{x})=t^2-x_1^2-x_2^2-x_3^2, \tag{4}$$

对应的真洛伦茨群. 我们将在下面详细地说明它的实质. 对于1维情形的洛伦茨变换, 我们已经得到了一个显形式的公式(3). 一般地保持$q(\mathbf{x})$ 的变换的公式看起来相当繁杂. 为此, 我们选择群$\mathbf{L}$的另一种描述方式. 也就是研究所有的2阶埃尔米特矩阵

$$P_{\mathbf{x}}=\begin{pmatrix} t-x_3 & x_2-ix_1\\ x_2+ix_1 & t+x_3\end{pmatrix}. \tag{5}$$

构成的空间. 这里, $\mathbf{x} = (t, x_1, x_2, x_3)$ 是4维实空间$\mathbb{R}^4$ 的向量. 向量与埃尔米特矩阵之间的对应

$$P_{\alpha\mathbf{x}+\beta\mathbf{y}} = \alpha P_{\mathbf{x}} + \beta P_{\mathbf{y}}$$

是相互单值的, 而且是线性的. 每个复矩阵

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad \alpha\delta - \beta\gamma = 1,$$

也就是群$SL_2(\mathbb{C})$ 的每个元素提供一个埃尔米特矩阵空间上的变换Γ_A ,

$$\Gamma_A(P_{\mathbf{x}}) = A \cdot P_{\mathbf{x}} \cdot A^*.$$

容易看出

$$\{\Gamma_A(P_{\mathbf{x}})\}^* = A^{**} P_{\mathbf{x}}^* A^* = \Gamma_A(P_{\mathbf{x}}),$$

(其中$A^* = {}^t\bar{A}$)就是通常的埃尔米特共轭. 因为,

$$\Gamma_A\left(\Gamma_B(P_{\mathbf{x}})\right) = ABP_{\mathbf{x}}B^*A^* = ABP_{\mathbf{x}}(AB)^* = \Gamma_{AB}(P_{\mathbf{x}}),$$

所以,

$$\Gamma_A\Gamma_B = \Gamma_{AB},$$

而且, 算子Γ_A 还是线性的

$$\Gamma_A(\alpha P_{\mathbf{x}} + \beta P_{\mathbf{y}}) = \alpha\Gamma_A(P_{\mathbf{x}}) + \beta\Gamma_A(P_{\mathbf{y}}).$$

反过来注意到

$$\det A \cdot \det P_{\mathbf{x}} \cdot \det A^* = \det P_{\mathbf{x}},$$

这是因为, 按条件$\det\ A = \det A^* = 1$. 而

$$\det P_{\mathbf{x}} = t^2 - x_1^2 - x_2^2 - x_3^2.$$

这就意味着, Γ_A 不改变二次型(4). 特别地, $\det\ \Gamma_A = \pm 1$. 事实上, 根据拓扑学的简单结果(函数det 的连续性和群$SL_2(\mathbb{C})$ 的连通性)可以推出$\det\ \Gamma_A = +1$. 我们可以对此事就信以为真, 尽管不需要太大的努力就可以证实这一点.

方程式

$$t^2 - x_1^2 - x_2^2 - x_3^2 = 0 \tag{6}$$

在$\mathbb{R}^4$ 中决定了一个锥面——一个特殊的2阶曲面(或者, 如同我们下一章要讲的, 二次曲面). 在这个曲面上, 包含经过坐标原点且经过曲面上任意一点的整个直线. 由条件$t > 0$ 得出一个锥面(6)的所谓的**上部空腔**.

其次, 不等式

$$t > 0, \quad t^2 - x_1^2 - x_2^2 - x_3^2 > 0$$

给出了矩阵$P_{\mathbf{x}}$ 正定的充要条件, 或者, 同样的, 对应的二次型的正定(见第1章§4的第8目)的充要条件.

显然, 这些正定性条件对矩阵

$$\Gamma_A(P_{\mathbf{x}}) = A \cdot P_{\mathbf{x}} \cdot A^*$$

也是保持不变的. 这意味着, 线性算子Γ_A 不仅保持锥面(6)不变, 而且也保持它的上部空腔不变.

Γ_A 的性质可以被概括成

1) Γ_A 是二次型(4)的自同构;

2) $\det\ \Gamma_A = 1$;

3) Γ_A 保持锥面(6)的上部空腔不变.

定义2 所有满足条件1)—3)的$\mathbb{R}^4 \to \mathbb{R}^4$ 的线性算子均被称为**真洛伦茨变换**, 而所有的这种变换构成的群$\mathbf{L}^+$ 就被称为**真洛伦茨群**.

实际上, 同态映射Γ 是个满同态(见习题3). 现在来计算核$\mathrm{Ker}\ \Gamma$. 设, 对任意埃尔米特矩阵$P_{\mathbf{x}}$ 有

$$P_{\mathbf{x}} = A \cdot P_{\mathbf{x}} \cdot A^*$$

(条件$\Gamma_A = \mathcal{E}$). 特别地, 当$\mathbf{e} = (1\ ,\ 0\ ,\ 0\ ,\ 0)$ 时, 我们有$P_{\mathbf{e}} = E$ 且$AA^* = E$, 从而$A^* = A^{-1}$. 这样一来, 就有

$$AP_{\mathbf{x}} = P_{\mathbf{x}}A.$$

选取不同的无关的矩阵$P_{\mathbf{x}}$, 就可以得到$A = \alpha E$. 又因为$\det\ A = 1$, 所以$\alpha = \pm 1$, 由此可见, $\mathrm{Ker}\ \Gamma = \{\pm E\}$.

我们就得到了如下的论断.

定理1 在行列式值为1的2阶复矩阵与固有的洛伦茨变换之间的对应$\Gamma : A \mapsto \Gamma_A$ 是群$SL_2(\mathbb{C})$ 到所有的真洛伦茨变换构成的群$\mathbf{L}^+$ 的一个同态映射. 每个真洛伦茨变换恰由两个复矩阵A和$-A$对应, 它们只差一个负号.

鉴于定理1, 常常把$SL_2(\mathbb{C})$ 称为洛伦茨群, 尽管更准确的说法应该是它的商群$SL_2(\mathbb{C})/\{\pm E\}$ 才对.

因为, 按定义, 二次型$q(\mathbf{x})$ 对于洛伦茨变换是不变的, 所以这个变换把由方程式

$$t^2 - x_1^2 - x_2^2 - x_3^2 = c, \quad c \in \mathbb{R}.$$

给出的曲面S_c转化成自己. 如果$c > 0$, 则S_c 是个双叶双曲面, 如果$c < 0$, S_c 就是一个单叶双曲面; 最后, S_0 是个锥面(这个从解析几何学中3维空间借用来的术语, 在下一章我们会频繁地使用). 在这些曲面中的每一个上都是具有同样意义的运动, 也

就是$\mathbb{R}^n$ 上的正交算子决定的球面S^{n-1} 上的运动(所谓运动就是一个保持点之间距离的变换).

双叶双曲面

$$t^2 - x_1^2 - x_2^2 - x_3^2 = 1, \quad t > 0$$

的上部空腔连同在其上决定的运动群$\mathbf{L}^+$ (或者$SL_2(\mathbb{C})$)可以表示成罗巴切夫斯基空间Λ^3 的模型之一. 我们目前不想在罗巴切夫斯基空间的概念上作更充分地停留, 而把注意力转向一种情况. 我们只对那样一种空间S的运动群G感兴趣, 对任意点$\dot{p} \in S$,都可以经过一个运动$g \in G$ 变到任意其他的点$\dot{q} \in S$, $g(\dot{p}) = \dot{q}$, 或者, 等价地, 任意点$\dot{q} \in S$ 都相对于某个$g \in G$ 而言是某一个点$\dot{p}_0$ 的像. 我们在§3中已经注意到了在仿射空间$\mathbb{A}$上群$\mathrm{Aff}(\mathbb{A})$ 的传递性, 而$\mathrm{Iso}(\mathbb{E})$ 在欧几里得空间上的作用也有传递性. 群$O(n)$ 作用在球面$S^{n-1} \subset \mathbb{R}^n$ 上, 显然, 也是可传递的(怎样简要地证明这些说法).

现在, 我们说, **洛伦茨群$\mathbf{L}^+$ 在Λ^3 上是传递的**. 对于点$\mathbf{x} = (t\ ,\ x_1\ ,\ \ x_2\ ,\ \ x_3\ ) \in \Lambda^3$, 如同过去一样, 我们把它和埃尔米特矩阵$P_{\mathbf{x}}$ 一起相提并论(见(5)). 正如我们已经知道的, 任意一个这样的矩阵都可以表示成

$$P_{\mathbf{x}} = A \cdot A^* = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^*$$

的样子, 其中$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ 是一个行列式为1的复矩阵. 这就意味着, 借助于运动Γ_A, 由固定的矩阵$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 可以得到$P_{\mathbf{x}}$.

点$\mathbf{x}_0 = (1,0,0,0)$的稳定子群$\mathbf{L}^+_{\mathbf{x}_0}$. 可以作为所有满足条件

$$A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

也就是$AA^* = E$ 的Γ_A 的集合. 又因为$\det\ A = 1$, 所以, 我们得到结论

$$\mathbf{L}^+_{\mathbf{x}_0} \cong SU(2)/\{\pm E\} \cong SO(3)$$

(最后面的那个同构将在[BA Ⅲ]中建立. 目前, 它对我们来说还不是必需的). 空间Λ^3 的运动还被称为**双曲面旋转**.

习　　题

1. 更充分地讨论一下算子$\mathcal{F} \in \mathbf{L}_1$ 作用的第二种可能性.
2. 证明, 对于第4目中定义的线性算子Γ_A , 行列式$\det\ \Gamma_A = 1$.

3. 证明, 同态映射

$$\Gamma : SL_2(\mathbb{C}) \to \mathbf{L}^+$$

实际上是个满射, 也就是一个到整个群$\mathbf{L}^+$ 上的映射.

4. 花费一些时间阅读教学参考书[2]中的§12, 以加强自己对素有的闵可夫斯基空间和洛伦茨群的物理观点.

第5章 二次曲面

在这一章里, 我们将要在仿射空间和欧几里得空间中研究读者在解析几何教程中已经熟悉了的几何图形, 用术语来说就是曲线, 曲面, 等等. 但是, 在前一章的结尾, 我们已经看到, 突破3维空间的限制是很有必要的. 高维空间中2阶曲面的彻底分类并不非常复杂: 在我们的安排之下形成了适当的分支结构, 尽管直观的几何想象能力大概会退居到次要地位. 主要的是用投影的观点来研究几何对象.

§1 二次函数

1. 仿射空间上的二次函数 设$\mathbb{A}$是域$\mathfrak{K}$上与向量空间V相伴随的一个m维仿射空间. 我们把$\mathfrak{K}$视为一个1维的仿射空间$\mathfrak{K}_a$. 例如, 实数集合$\mathbb{R}$同时也是一条实的仿射直线$\mathbb{R}_a = \mathbb{R}$, 1维的仿射群就可以作用其上(见第3章§1和§3).

类似于双线性函数$f : V \times V \to \mathfrak{K}$, 也可以定义双仿射函数.

定义1 称函数$\Phi : \mathbb{A} \times \mathbb{A} \to \mathfrak{K}$是**双仿射的**, 如果$\Phi(\dot{p}, \dot{q})$在固定点$\dot{p}$或固定点$\dot{q}$时, 它就是一个仿射映射$\dot{q} \mapsto \alpha \in \mathfrak{K}_a$(或者$\dot{p} \mapsto \alpha \in \mathfrak{K}_a$). 称双仿射函数是**对称的**. 如果

$$\Phi(\dot{p}, \dot{q}) = \Phi(\dot{q}, \dot{p}) \ \ \forall \dot{p}, \dot{q} \in \mathbb{A}.$$

我们不去证明这样一个事实, 如果选择某一个点$\dot{o} \in \mathbb{A}$作为“原点”并设$\dot{p} = \dot{o} + \mathbf{x}, \dot{q} = \dot{o} + \mathbf{y}$, 那么, 任意双仿射函数$\Phi$都可以表达成

$$\Phi(\dot{o} + \mathbf{x}, \dot{o} + \mathbf{y}) = f(\mathbf{x}, \mathbf{y}) + l(\mathbf{x}) + l'(\mathbf{y}) + \varphi_0 \tag{1}$$

的形式, 其中f是V上的双线性函数, 而$\varphi_0 = \Phi(\dot{o}, \dot{o})$是个纯量. 固定向量$\mathbf{y} = \mathbf{a}$且令

$$h(\dot{o}) = l'(\mathbf{a}) + \varphi_0, \quad Dh(\mathbf{x}) = f(\mathbf{x}, \mathbf{a}) + l(\mathbf{x}),$$

我们不难相信, 实际上的变换

$$\dot{p}=\dot{o}+\mathbf{x}\mapsto h(\dot{o}+\mathbf{x})=h(\dot{o})+Dh(\mathbf{x})$$

是仿射–线性的, 它以函数$Dh:V\to\mathfrak{K}$为其线性部分.

为简便起见, 根据上面说的那些充分理由, 我们可以把(1)当作双仿射函数的定义. 很容易相信, 对称的双仿射函数Φ可以写成

$$\Phi(\dot{o}+\mathbf{x},\dot{o}+\mathbf{y})=f(\mathbf{x},\mathbf{y})+l(\mathbf{x})+l(\mathbf{y})+\varphi_0, \tag{2}$$

它带有对称的双线性函数f: $V\times V\to\mathfrak{K}$, 线性函数$l:V\to\mathfrak{K}(l'=l)$以及一个纯量$\varphi_0=\Phi(\dot{o},\dot{o})\in\mathfrak{K}$.

定义2 设$Q(\dot{p})=\Phi(\dot{p},\dot{p})$, 其中$\Phi$形如(2), 我们就称$Q:\mathbb{A}\to\mathfrak{K}$是$\mathbb{A}$上的一个**二次函数**.

设q是V上的二次型: $q(\mathbf{x})=f(\mathbf{x},\mathbf{x})$. 与(2)对应, 应有

$$Q(\dot{o}+\mathbf{x})=q(\mathbf{x})+2l(\mathbf{x})+\varphi_0 \tag{3}$$

在$\mathbb{A}$中取一个以$\dot{o}$为原点的坐标系$\{\dot{o};\mathbf{e}_1,\cdots,\mathbf{e}_n\}$, 我们就得到二次函数$Q$在点$\dot{p}=\dot{o}+\mathbf{x},\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n$的值的坐标记法

$$Q(\dot{o}+\mathbf{x})=\sum_{i,j=1}^{n}\varphi_{ij}x_ix_j+2\sum_{i=1}^{n}\varphi_ix_i+\varphi_0, \tag{4}$$

在这里, 系数φ_{ij}组成一个对称矩阵$F=(\varphi_{ij})$. 比较(3)和(4), 我们发现, 在其他坐标系之下也可以用同样的方式来刻画$Q(\dot{p})$, 即写成次数$\leqslant 2$ 的多项式形式, 尽管它们是另外的一些系数.

设$x'_1,\cdots,x'_n$是点$\dot{p}=\dot{o}'+\mathbf{x}'$在坐标系$\{\dot{o}';\ \mathbf{e}'_1,\cdots,\mathbf{e}'_n\}$之下的坐标. 正如我们已经知道的(参见第4章§1的(3)), 原来的坐标可以用新的坐标表示出来

$$x_i=\sum_{i=1}^{n}a_{ij}x'_j+b_i,\qquad i=1,\cdots,n,$$

这里$A=(a_{ij})$是个非退化矩阵. 二次型q在新的坐标系之下的矩阵是${}^tA\cdot F\cdot A$(见第1章§4的(5)式). 特别地, 矩阵的秩相对于仿射变换是不变的, 从而可令

$$\operatorname{rank}Q:=\operatorname{rank}q=r.$$

2. 二次函数的中心点 现在, 还可以引进一个有用的概念. 为此, 我们研究Q在点$\dot{q}=\dot{p}+\mathbf{y}=\dot{o}+\mathbf{x}+\mathbf{y}\in\mathbb{A}$处的值. 由(3)可以直接得出

$$Q(\dot{q})=Q(\dot{p}+\mathbf{y})=Q(\dot{o}+\mathbf{x}+\mathbf{y})=q(\mathbf{x}+\mathbf{y})+2l(\mathbf{x}+\mathbf{y})+\varphi_0,$$

也就是

$$Q(\dot{q}) = Q(\dot{p}) + q(\mathbf{y}) + 2\{f(\mathbf{x},\mathbf{y}) + l(\mathbf{y})\} \tag{5}$$

(记住, f是与q配极的双称的双线性型).

定义3 称点$\dot{p} \in \mathbb{A}$是一个关于二次函数Q的**中心**(或者, 一个**中心的点**), 如果

$$Q(\dot{p} + \mathbf{y}) = Q(\dot{p}) + q(\mathbf{y}), \quad \forall \mathbf{y} \in V. \tag{6}$$

用$C(Q)$代表二次函数Q的所有的中心的集合. 对于$C(Q) \neq \varnothing$的二次函数Q, 就称它是**中心的**.

比较(5)和(6)即可表明, 条件$\dot{p} = \dot{o} + \mathbf{x} \in C(Q)$可以写成

$$f(\mathbf{x},\mathbf{y}) + l(\mathbf{y}) = 0, \quad \forall \mathbf{y} \in V. \tag{7}$$

特别地, 当公式(3)中的线性函数l为零时, 零$\dot{o}$就是中心的. 换句话说, 如果坐标原点$\dot{o}$是个中心点, 那么, $Q(\dot{o} + \mathbf{x})$的表达式就不含$x_1, \cdots, x_n$的一次项. 对于$\dot{o} \in C(Q)$, 点$\dot{o}' = \dot{o} + \mathbf{b}$为中心点的条件(7)必形如$f(\mathbf{b},\mathbf{y}) = 0$. 这意味着$q(\mathbf{b}) = f(\mathbf{b},\mathbf{b}) = 0$. 回顾$\varphi_0 = Q(\dot{o})$, 我们就由(3)得到等式$Q(\dot{o}') = Q(\dot{o})$. 这样一来, 就有

$$\dot{o}, \dot{o}' \in C(Q) \Rightarrow Q(\dot{o}) = Q(\dot{o}'). \tag{8}$$

怎样辨认一个由公式(3)或公式(4)给出的二次函数Q是不是中心的呢? 进一步, 如果它是中心的, 又如何找出$C(Q)$呢?

为了给出这些问题的答案, 应当从条件(7)出发, 等价地, 也容易看出, 可以从方程组

$$f(\mathbf{e}_i,\mathbf{x}) + l(\mathbf{e}_i) = 0, \qquad i = 1, 2, \cdots, n$$

出发. 于是, 对于中心点$\dot{p} = \dot{o} + \mathbf{x}$决定的向量$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$的坐标, 就得到方程组

$$\sum_{j=1}^{n} \varphi_{ij} x_j = -\varphi_i \qquad i = 1, \cdots, n, \tag{9}$$

借助克罗内克–卡佩利定理可以确信该方程组的相容性. 在方程组(9)相容的情况下, 集合$C(Q)$或者是一个点($\det F \neq 0$情形), 或者, 由刚刚提到的那个定理推出, 它是个$n-r$维的仿射子空间. 这个空间的方向平面U与线性方程组

$$\sum_{j=1}^{n} \varphi_{ij} x_j = 0, \qquad i = 1, \cdots, n,$$

的解空间重合. 从而

$$U = \operatorname{Ker} q = \operatorname{Ker} f$$

(回忆一下, $\operatorname{Ker} q = \{\mathbf{x} \in V \mid f(\mathbf{x},\mathbf{y}) = 0, \forall \mathbf{y} \in V\}$).

我们就证明了

定理1 对应于(4), 在坐标系$\{\dot{o};\mathbf{e}_1,\cdots,\mathbf{e}_n\}$之下, 给定的二次函数$Q$的中心集$C(Q)$由线性方程组(9)决定的点$\dot{p}=\dot{o}+\mathbf{x}$组成. 如果$\dot{o}'=\dot{o}+\mathbf{x}'$是一个中心点, 那么

$$C(Q)=\dot{o}'+U$$

刚好是以$U=\operatorname{Ker}q$为方向平面的仿射空间. 其次, $C(Q)=\varnothing\Rightarrow r<n$; $C(Q)$是仿射不变的, 它只由函数Q决定.

3. 把二次函数化成规范型 如果存在仿射自同构$g\in\operatorname{Aff}(\mathbb{A})$使得$Q'=Q\cdot g$, 就称$\mathbb{A}$上的两个二次函数$Q$和$Q'$是在$\mathbb{A}$上**仿射等价的**.

对于$Q(\dot{p})$的最简单的表达的愿望, 可以由下面的定理来实现.

定理2 设Q是域$\mathfrak{K}$上的n维仿射空间$\mathbb{A}$上的一个秩为r的二次函数. 如果$C(Q)$是个空集, 当然, 这也意味着$r<n$, 那么, 借助适当地选择坐标系$\{\dot{o};\mathbf{e}_1,\cdots,\mathbf{e}_n\}$, 二次函数$Q$可以化成

$$Q(\dot{o}+\mathbf{x})=\alpha_1x_1^2+\cdots+\alpha_rx_r^2+2x_{r+1}, \tag{10}$$

其中$\alpha_1,\cdots,\alpha_r$是纯量; 这种情况下, Ker q就是方程组$x_1=\cdots=x_r=0$的解空间(q是与Q相关联的二次型).

如果Q是中心的, 那么, 选择适当的中心点$\dot{o}$为原点的坐标系, 就可将其化成

$$Q(\dot{o}+\mathbf{x})=\alpha_1x_1^2+\cdots+\alpha_rx_r^2+\varphi_0; \tag{11}$$

在这种情形, 对任意点$\dot{o}'\in C(Q)$都有$Q(\dot{o}')=\varphi_0$. 而且(10)型和(11)型的二次函数不是仿射等价的.

证明 定理本身要比它的叙述简单些. 开始, 在V中选择一个对于二次型q的规范基底(见第1章§4). 在对应的坐标系$\{\dot{o}';\mathbf{e}'_1,\cdots,\mathbf{e}'_n\}$之下, 函数$q$形如

$$Q(\dot{o}'+\mathbf{x})=\alpha_1x'^2_1+\cdots+\alpha_rx'^2_r+2\beta'_1x'_1+\cdots+2\beta'_nx'_n+\gamma'$$

其中$\alpha_1\neq0,\cdots,\alpha_r\neq0$. 把原点移到$\dot{o}''$, 导出来坐标变换是

$$\begin{aligned}x''_1&=x'_i+\beta'_i/\alpha_i,\quad i=1,\cdots,r,\\x''_i&=x'_i,\qquad i=r+1,\cdots,n,\end{aligned}$$

得到在线性部分$x''_1,\cdots,x''_r$前面的系数均为零, 所以

$$Q(\dot{o}'')=\alpha_1x''^2_1+\cdots+\alpha_rx''^2_r+2\beta''_{r+k}x''_{r+k}+\cdots+2\beta''_nx''_n+\gamma''.$$

如果不是所有的β''_j都等于零(前面已设$\beta''_{r+k}\neq0$), 那么, 再一次做坐标的仿射替换

$$\begin{aligned}x_i&=x''_i,\qquad i=1,\cdots,r,\\x_{r+1}&=\beta''_{r+k}x''_{r+k}+\cdots+\beta''_nx''_n+\gamma''/2,\end{aligned}$$

$$x_{r+2}=x''_{r+1},\quad \cdots,\quad x_{r+k}=x''_{r+k-1},$$
$$x_{r+k+i}=x''_{r+k+i},\qquad i\geqslant 1,$$

就把Q化成了如(10)的形式. 在Q的另外一种相反的情形, 用准确的表达方式就获得如(11)的形式.

换言之, 我们可以认为Q已经化成了如(10)或者(11)的形式. 因为$q(\mathbf{x})=\alpha_1x_1^2+\cdots+\alpha_rx_r^2$,所以, Ker q是V的一个由方程组$x_1=\cdots=x_r=0$给出的$n-r$维的子空间. 设$\dot p=\dot o+\mathbf{x}$是个中心点, $\dot p+\mathbf{y}$是任意点, 那么, 对型(10)有

$$\begin{aligned}Q(\dot p+\mathbf{y})=Q(\dot o+\mathbf{x}+\mathbf{y})&=\sum_i\alpha_i(x_i+y_i)^2+2(x_{r+1}+y_{r+1})\\&=Q(\dot p)+q(\mathbf{y})+2\sum_{i=1}^{r}\alpha_ix_iy_i,\end{aligned}$$

而对型(11)有

$$Q(\dot p+\mathbf{y})=Q(\dot p)+q(\mathbf{y})+2\sum_{i=1}^{r}\alpha_ix_iy_i.$$

点$\dot p$的中心性条件(6)需要对任意的$\mathbf{y}\in V$都满足, 在刚刚说过的后一种情况下可归纳为$x_1,\cdots,x_r=0$, 也就是归结为$\mathbf{x}\in\operatorname{Ker}q$. 而第一种情形, 由于存在自由项$2y_{r+1}=0$, 它一般是不能被满足的, 从而$C(Q)=\varnothing$. □

推论 *在实数域$\mathbb{R}$上, 每个二次函数Q借助于适当地选择$\mathbb{A}$的坐标系$\{\dot o;\mathbf{e}_1,\cdots,\mathbf{e}_n\}$都可以简化成规范型*

$$Q(\dot o+\mathbf{x})=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+2x_{r+1},\tag{12}$$

$$Q(\dot o+\mathbf{x})=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2+\varphi_0.\tag{13}$$

而且是唯一确定的.

证明 因为与Q相关联的二次型q的正惯性指数s以及秩r对于非退化线性变换都是不变的(见第1章§4的定理5), 所以, 定理2就给出了所需的一切. □

可以用另外的叙述方式补充一些推论.

$\mathbb{A}$上两个二次函数Q, Q'仿射等价, 当且仅当, 它们具有相同的秩和相同的符号差, 并且, 或者它们双双都是非中心的, 或者同时都是中心的且在相应的中心点上取得同一个值.

4. 欧几里得空间上的二次函数 在(点的)欧几里得空间$(\mathbb{E},V)$的情形, 很自然地要研究相对于自同构群$\operatorname{Iso}(\mathbb{E})$作用的二次函数的等价性.

定义4 *称$\mathbb{E}$上两个二次函数Q_1,Q_2是*Iso($\mathbb{E}$)**等价**的, *如果存在一个运动$g\in\operatorname{Iso}(\mathbb{E})$使得$Q_2=Q_1\cdot g$, 也就是$Q_2(\dot p)=Q_1(g(\dot p))$.*

定理3 *n维空间$\mathbb{E}$上的任意二次函数Q都可以借助在$\mathbb{E}$中选择适当的直角坐标系$\{\dot o;\mathbf{e}_1,\cdots,\mathbf{e}_n\}$简化成下列两种形式之一:*

$$Q(\dot o+\mathbf{x})=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+\varphi_0,\qquad \dot o\in C(Q),\tag{14}$$

$$Q(\dot{o}+\mathbf{x})=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+2\mu x_{r+1},\qquad \mu>0. \tag{15}$$

所有的实数λ_i都不等于零, 上述形式, 不计变量x_i的变号顺序, 是唯一确定的.

证明 设$Q=(\dot{o}'+\mathbf{x})=q(\mathbf{x})+2l'(\mathbf{x})+Q(\dot{o}')$. 我们从在欧几里得向量空间$V$中选择标准正交基$(\mathbf{e}_1',\cdots,\mathbf{e}_n')$开始, 让$q$在这个基底之下具有规范形式

$$q(\mathbf{y})=\sum_{i=1}^{r}\lambda_iy_i^2,\quad \lambda_i\neq 0,\quad \mathbf{y}=\sum_{i=1}^{n}y_i\mathbf{e}_i'.$$

这种选择可由第3章§3的定理7给予保障. 如果Q是中心函数, 那么, 在必要的情形, 用$\dot{o}'$代替$\dot{o}$(可如同在仿射情形一样地找到), 我们可以导出表达式(14). 在中心函数的情形, 我们有

$$Q(\dot{o}')+\mathbf{y}=\sum_{i=1}^{r}\lambda_iy_i^2+2\sum_{i=1}^{n}\varphi_i'y_i+\varphi_0'.$$

先完成坐标变换(移动原点)

$$\begin{aligned}&z_i=y_i+\varphi_i'/\lambda_i,\quad i=1,2,\cdots,r,\\&z_i=y_i,\qquad\qquad i=r+1,\cdots,n,\end{aligned}$$

我们就可以得到

$$Q(\dot{o}+\mathbf{z})=\sum_{i=1}^{r}\lambda_iz_i^2+2\sum_{i=r+1}^{n}\mu_iz_i+2\mu_0,$$

这里, 由于Q的中心性, 对于$i>r$, 不能所有的μ_i都等于零. 我们引入线性函数$\sum\limits_{i=r+1}^{n}\mu_iz_i$的“模”:

$$\mu=\sqrt{\mu_{r+1}^2+\cdots+\mu_n^2}>0,$$

从而可得坐标变换

$$\begin{aligned}&x_i=z_i,\quad i=1,2,\cdots,r,\\&x_{r+1}=\sum_{k=r+1}^{n}\frac{\mu_k}{\mu}z_k+\frac{\mu_0}{\mu},\\&x_i=\sum_{j=r+1}^{n}\alpha_{ij}z_j,\qquad i=r+2,\cdots,n,\end{aligned}$$

它可以由$(n-r)\times(n-r)$阶矩阵A

$$A=\begin{pmatrix}\mu_{r+1}/\mu & \mu_{r+2}/\mu & \cdots & \mu_n/\mu\\ \alpha_{r+2,r+1} & \alpha_{r+2,r+2} & \cdots & \alpha_{r+2,n}\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{n,r+1} & \alpha_{n,r+2} & \cdots & \alpha_{n,n}\end{pmatrix}$$

来实现.

因为我们只想利用直角坐标架, 所以矩阵A需要选择成正交矩阵. 它的第一列元素的平方和为1, 根据假定, 元素a_{ij}可由我们安排. 所以A是可以构作出来的(见第1章§3的第5目中与此相关的内容). 替换之后, 就有

$$Q(\dot{o}+\mathbf{x})=\sum_{i=1}^{r}\lambda_i x_i^2+2\mu x_{r+1},\qquad \mathbf{x}=\sum_{i=1}^{n}x_i\mathbf{e}_i,$$

这正是公式(15)所要求的.

我们用下面的想法来证明型(14)和型(15)的唯一性. 依据关于化秩为r的二次型q到主轴上去的定理, 数值λ_i是唯一确定的. 数$\varphi_0 = Q(\dot{o})$与中心$\dot{o}$的选择没有关系(见(8)).

剩下来, 我们还需要证明, 常数$\mu>0$的选择与个人意愿没有关系. 假设在某个直角坐标系$\{\dot{o}\,;\mathbf{e}_1',\cdots,\mathbf{e}_n'\}$中, 有

$$Q(\dot{o}'+\mathbf{x})=\sum_{i=1}^{r}\lambda_i (x_i')^2+2\mu' x_{r+1}',\qquad \mu'>0.$$

设$\mathcal{F}$是V上的对称的线性函数, 对应的双线性函数是f, 与q配极(见第3章§3的第1目)

$$f(\mathbf{x},\mathbf{y})=(\mathcal{F}\mathbf{x}|\mathbf{y}).$$

它的矩阵

$$F=\mathrm{diag}(\lambda_1,\cdots,\lambda_r,0,\cdots,0)$$

在基底$(\mathbf{e}_i)$之下和在基底$(\mathbf{e}_i')$之下具有同样的形式. 这意味着

$$\mathrm{Im}\,\mathcal{F}=\langle \mathbf{e}_1,\cdots,\mathbf{e}_r\rangle=\langle \mathbf{e}_1',\cdots,\mathbf{e}_r'\rangle,$$

因此, 由$(\mathbf{e}_i)$向$(\mathbf{e}_i')$的转换矩阵形如

$$B=\begin{pmatrix}B_1 & 0\\ 0 & B_2\end{pmatrix}$$

其中B_1是$r\times r$阶的正交矩阵, B_2是个$(n-r)\times(n-r)$阶的正交矩阵. 计算坐标原点的位移和Q的表达式中的指数$> r+1$的坐标的缺项, 就得到

$$\sum_{i=1}^{r}\lambda_i x_i^2=\sum_{i=1}^{r}\lambda_i (x_i')^2+2\nu,$$

$$2\mu x_{r+1}=2\mu' x_{r+1}'-2\nu,\quad \nu\in\mathbb{R}.$$

从而

$$x_{r+1}=\frac{\mu'}{\mu}\cdot x_{r+1}'-\frac{\nu}{\mu}.$$

由于B的正交性, 在表达式$x_{r+1}=\sum\alpha_j x'_j+\nu_0$中应当满足等式$\sum\alpha_j^2=1$, 在我们这种情形可导出关系式

$$(\mu'/\mu)^2=1.$$

从而就得到了我们所需要的$\mu'=\mu$, 因为μ'和μ都是正的. □

习　　题

1. 假定$2s\geqslant r>0$, 找出在n维的实仿射空间上中心的二次函数的等价类的个数.

2. 找出在$\mathbb{R}$上形如

$$x_1^2+2\sum_{1\leqslant i<j\leqslant n}x_ix_j+2\sum_{i=1}^{n}x_i+1$$

的二次函数Q的中心集$C(Q)$.

§2　仿射空间与欧几里得空间中的二次曲面

1. 二次曲面的一般概念　$\mathbb{A}$上的每个二次函数Q都可以建立一个与点S_Q的空间图形的对应. 这个图形被称为**二次曲面**(或**二阶曲面**(**超曲面**)). 它的定义是满足方程式$Q(\dot{p})=0$的所有点的几何位置(集合). 当$n=2$时, 也称二次曲面为**圆锥曲线**(**二阶曲线**). 可以研究任意域$\mathcal{K}$上的二次曲面(实际上, 可能会在不同的问题上遇到它们), 但是, 最自然的是把$\mathcal{K}$取成代数封闭域$\mathbb{C}$. 然而, 出于理解的直观性(同样也出于某些条件下提出的问题), 我们限制在$\mathcal{K}=\mathbb{R}$条件下. 刚开始, 我们暂且把"零"二次曲面排除在研究范围之外, 这样会更方便一些, "零"二次曲面就是没有任何一点位于其上的二次曲面. 比方说, 二次函数$x_1^2+x_2^2+1$就定义了一个零曲线. 更准确地说: 在将来, 总是假定, 由方程式$Q(\dot{p})=0$决定的二次曲面S_Q是个非空集合, 且

$$\operatorname{rank}S_Q:=r=\operatorname{rank}Q=\operatorname{rank}q>0,$$

我们同样默认, $n\geqslant 2$.

定义1　称二次曲面是$\mathbb{A}$的**二重子空间**, 如果它与另一个仿射平面重合.

例如, n维空间$\mathbb{A}$上的方程$x_1^2+\cdots+x_r^2=0$等价于方程组$x_1=0,\cdots,x_r=0$, 可见, 决定了一个$n-r$维的子空间. 二重子空间的定义与坐标系没有关系, 因此, 决定S_Q的二次函数Q可以取成规范型. 正如在定理2的推论中推出的一样, 任意二重子空间都可以由前面研究过的形如$x_1^2+\cdots+x_r^2=0$的方程给定. 需要注意, 二重(线性)子空间$x_1^2+x_2^2=0$与二重子空间$2x_1^2+3x_2^2=0$在3维空间中对应同一条直线$x_1=0,x_2=0$. 情况将要完全改变, 我们着手处理令人满意得多的不同于二重子空间的情形.

定理1(唯一性定理)　如果二次曲面S不是二重子空间, 那么, 它的任意两个方程式(在同一坐标系之下)必然是成比例的, 即

$$S_{Q_1}=S=S_{Q_2}\Rightarrow Q_2=\lambda Q_1,\quad \lambda\in\mathbb{R}^*.$$

证明 按照条件, 我们的二次曲面由两个方程式: $Q_1(\dot{p})$和$Q_2(\dot{p})$给出. 只需在§1的公式(12)和(13)上稍微看一下就足以相信S上不会少于两个不同的点. 进一步, 一定存在$\dot{p}$, $\dot{q} \in S$, 尽管经过它们的直线不一定整个地含在S中. 事实上, 若不然, 与第4章§1的定理4相对应, 二次曲面S就可以归结为仿射空间(平面), 也就是二重子空间. 容易看出, $\Pi_{\dot{p},\dot{q}} \cap S = \{\dot{p}, \dot{q}\}$是一个由两个点组成的集合.

我们以上述的性质固定这两点$\dot{p}$, $\dot{q} \in S$且选择$\dot{p}$为坐标原点, 而向量$\overrightarrow{\dot{p}\dot{q}}$作为向量空间$V$的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$的最后边的那个向量. 于是, $\Pi_{\dot{p},\dot{q}}$就是由坐标为$(0, \cdots, 0, \beta)$的点组成. 而点$\dot{p}$的坐标是$(0, \cdots, 0, 0)$, 点$\dot{q}$的坐标就是$(0, \cdots, 0, 1)$.

再把Q_1按坐标x_n的降幂写出来

$$Q_1(\dot{p} + \mathbf{x}) = \delta x_n^2 + g(x_1, \cdots, x_{n-1})x_n + h(x_1, \cdots, x_{n-1}).$$

这里, g是一次多项式, 而h是$x_1, \cdots, x_{n-1}$的二次多项式(g和h都不一定是线性多项式). 由于$\Pi_{\dot{p},\dot{q}}$与S相交于不同的两个点, 这意味着三项式

$$\delta x_n^2 + g(0)x_n + h(0)$$

有两个不同的实根, 也就是$g(0)^2 - 4\delta h(0) > 0$(事实上, $\delta \neq 0$, $g(0) \neq 0$, $h(0) = 0$). 对δ略作改动, 从一开始, 我们就可以假定$\delta = 1$. 同样的事实对Q_2也成立, 所以,

$$Q_i(\dot{p} + \mathbf{x}) = x_n^2 + g_i(x_1, \cdots, x_{n-1})x_n + h_i(x_1, \cdots, x_{n-1}), \qquad i = 1, 2,$$

而且$\Delta_1(0) > 0$, $\Delta_2(0) > 0$, 其中

$$\Delta_i(x_1, \cdots, x_{n-1}) = g_i(x_1, \cdots, x_{n-1})^2 - 4h_i(x_1, \cdots, x_{n-1}), \qquad i = 1, 2,$$

是Q_i关于变量x_n的判别多项式, 系数在$\mathbb{R}(x_1, \cdots, x_{n-1})$中.

按照我们规定的任务, 需要证明, $Q_2 = Q_1$. 我们任意选择$\lambda_1, \cdots, \lambda_{n-1} \in \mathbb{R}$后将其固定, 再来研究$\mathbb{A}$中的平面

$$x_1 = t\lambda_1, \quad \cdots, \quad x_{n-1} = t\lambda_{n-1}, \quad t \in \mathbb{R}. \tag{1}$$

那么, 对于

$$\mathbf{x} = t\lambda_1\mathbf{e}_1 + \cdots + t\lambda_{n-1}\mathbf{e}_{n-1} + x_n\mathbf{e}_n$$

将会有

$$Q_i(\dot{p} + \mathbf{x}) = x_n^2 + \tilde{g}_i(t)x_n + \tilde{h}_i(t), \tag{2}$$

其中

$$\tilde{g}_i(t) = g_i(t\lambda_1, \cdots, t\lambda_{n-1}), \quad \tilde{h}_i(t) = h_i(t\lambda_1, \cdots, t\lambda_{n-1}). \tag{3}$$

同样地, 我们设

$$\tilde{\Delta}_i(t) = \tilde{g}_i(t)^2 - 4\tilde{h}_i(t), \quad i = 1, 2.$$

按条件, $\tilde{\Delta}_1(0)>0$, $\tilde{\Delta}_2(0)>0$. 从而可以找到$\varepsilon>0$, 使得, 当$|t|<\varepsilon$时, 不等式

$$\tilde{\Delta}_1(t)>0, \qquad \tilde{\Delta}_2(t)>0$$

得到满足. 换一种说法就是对任意$|t|<\varepsilon$, 多项式(2)都有两个不同的实根. 但是, 根据条件, 这些多项式的根的集合对固定的t都是重合的——它就是S与子空间(1)的交集. 刚好二次标准多项式有唯一的一个根. 所以, 它们的系数是重合的

$$\tilde{g}_1(t)=\tilde{g}_2(t), \qquad \tilde{h}_1(t)=\tilde{h}_2(t), \qquad |t|<\varepsilon. \tag{4}$$

但是, 有无穷多个$t\in\mathbb{R}$, $|t|<\varepsilon$, 从而, 对所有的t, 等式(4)都成立. 特别地, 当$t=1$时可把(4)写成多项式函数形式

$$\begin{aligned} g_1(\lambda_1,\cdots,\lambda_{n-1})&=g_2(\lambda_1,\cdots,\lambda_{n-1}),\\ h_1(\lambda_1,\cdots,\lambda_{n-1})&=h_2(\lambda_1,\cdots,\lambda_{n-1}). \end{aligned} \tag{5}$$

由[BA I]我们知道, 两个次数都为m的多项式函数$f_i:\lambda\mapsto f_i(\lambda)$, 只要在$k\geqslant m+1$个不同的$\lambda$处相等, 那么, 作为多项式, 它们就是相同的. 即$f_1(X)=f_2(X)$. 推广到现在的多元多项式情形, 下列断言成立(见[BA I]第6章§1的习题2). 如果多项式$f_1(X_1,\cdots,X_{n-1})$和$f_2(X_1,\cdots,X_{n-1})$决定相同的多项式函数$\mathbb{R}^n\to\mathbb{R}$, 那么, 它们是相同的, 即它们的系数对应相等. 要证明这个断言, 只要把多项式按某一个变量的幂改写, 然后对n用归纳法.

利用这个论断, 我们可由(5)推出等式

$$\begin{aligned} g_1(x_1,\cdots,x_{n-1})&=g_2(x_1,\cdots,x_{n-1}),\\ h_1(x_1,\cdots,x_{n-1})&=h_2(x_1,\cdots,x_{n-1}), \end{aligned}$$

这说明, $Q_1=Q_2$. □

2. 二次曲面的中心 可以直接看出来, 图10上描述的二次曲面相对于坐标原点是对称的. 更一般的几何图形反映出

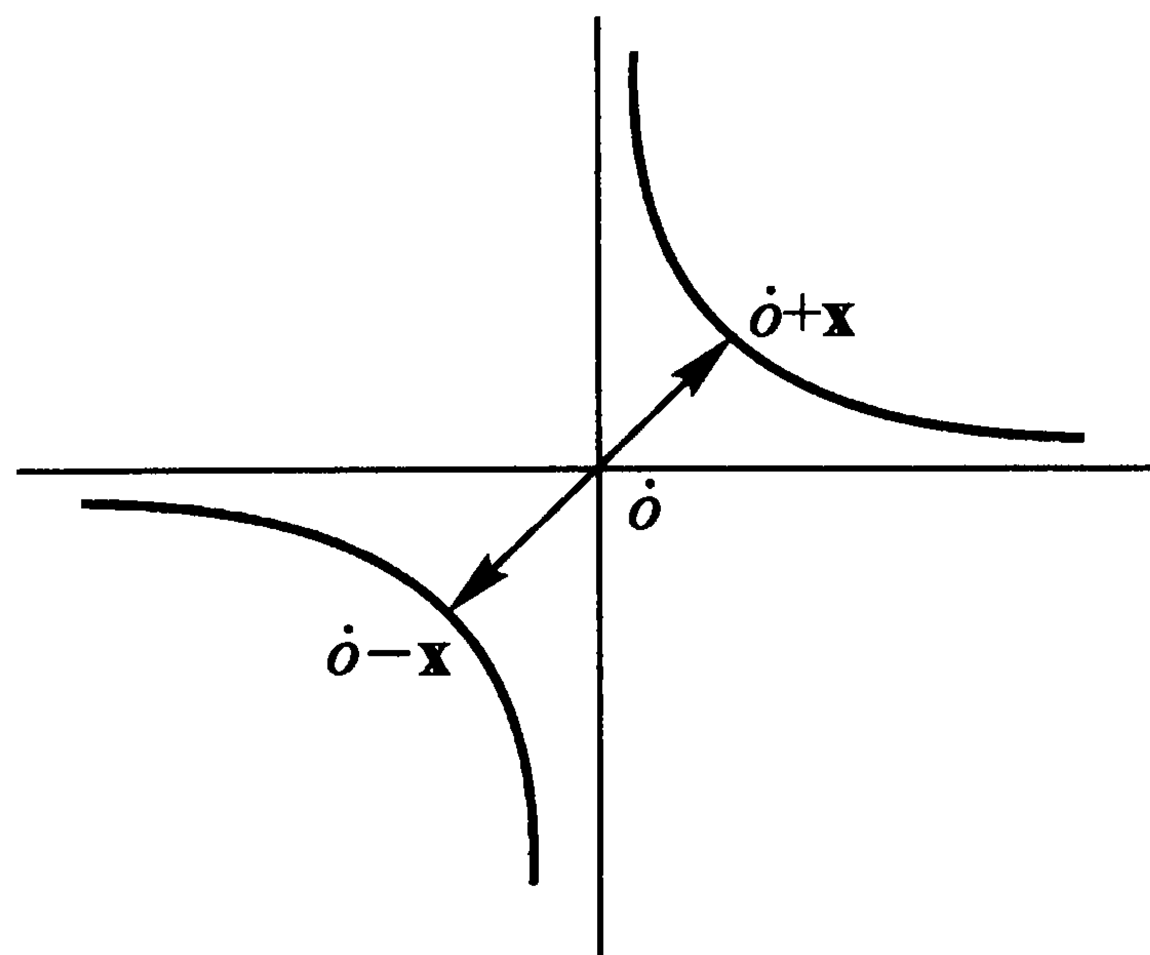


图10

定义2 称点$\dot{o}$为仿射空间$\mathbb{A}$中的二次曲面S_Q的**中心**(或**对称中心**), 如果对任意点$\dot{o}+\mathbf{x}$, 点$\dot{o}-\mathbf{x}$总是和它一起同时属于S_Q. 称二次曲面S是**有中心的**, 如果它有一个中心; 如果S没有中心, 就说它是**无中心的**.

假设中心在点$\dot{o}$的有中心的二次曲面S不是一个二重子空间. 再设

$$Q(\dot{o}+\mathbf{x})=q(\mathbf{x})+2l(\mathbf{x})+\varepsilon_0=0$$

是它的方程式. 由S的中心性, 二次函数$Q_1(\dot{o}+\mathbf{x}) := Q(\dot{o}-\mathbf{x})$确定了同样的二次曲面$S$:

$$Q_1(\dot{o}+\mathbf{x}) = q(\mathbf{x}) - 2l(\mathbf{x}) + \varepsilon_0 = 0.$$

根据定理1, 应有成比例的性质

$$Q_1 = \lambda Q, \quad \lambda \in \mathbb{R},$$

又因为$q \neq 0$, 所以, 这里只能有$\lambda = 1$且$l = 0$时才能成立. 但是, 由第2目, 我们已经知道$l = 0$正是点$\dot{o}$对于Q的中心性条件. 我们就推导出了这样的结论, 二次曲面(不是二重子空间)的中心与给出这个二次曲面的二次函数Q中心是重合的. 对称的二次曲面S_Q的中心的集合$C(S_Q)$与相对于二次函数Q的中心点的集合$C(Q)$是重合的, 而且(当非空时)是个仿射子空间(§1的定理1). 我们已经研究过它在任意坐标系下描述方式, 所以任意二次曲面S的中心性问题都可以获得有效解决.

3. 仿射空间中的二次曲面的规范型(典范型) 下面的定理是基本的

定理2 *在n维实的仿射空间中二次曲面的方程式经仿射自同构变换可以化成, 而且仅能化成下列标准型之一.*

在坐标原点处具有对称中心的可中心化的二次曲面情形穷尽以下类型:

$$I_{s,r}: x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 = 1, \qquad 0 < s \leqslant r;$$
$$I'_{s,r}: x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 = 0, \qquad r/2 \leqslant s \leqslant r.$$

无中心的二次曲面情形, 穷尽以下类型:

$$II_{s,r}: x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 = -2x_{r+1}, \qquad r/2 \leqslant s \leqslant r.$$

证明 几乎就是显然的: 只要利用§1定理2的推论并注意到对任意$\lambda \neq 0$都有$S_{\lambda Q} = S_Q$. 这就给出了, 在§1的表达式(13)中用-1替代φ_0(如果它不等于零)的可能性. 在$I_{s,r}$中条件$s > 0$排除了零二次曲面情形. 在$I'_{s,r}$中等式$s = r$对应二重子空间. □

定义3 *称$I_{n,n}$型的二次曲面为***椭球面**, *称$I_{s,n}$型$s < n$为***双曲面**, *$II_{n-1,n-1}$型为***椭圆抛物面**, *而$II_{s,n-1}$型为***双曲抛物面**, *所有这些曲面都是非退化的.*

当$r < n$时, $I_{s,r}$型和$I'_{s,r}$型以及$r < n-1$时的$II_{s,r}$型都被称之为**柱面**, 而将$I'_{s,n}$型称为**锥面**. 把锥面和柱面合起来称为**退化的二次曲面**.

锥面(图11)可以作为这样一种二次曲面S用不变的方式加以刻画, 即S上有一点$\dot{o}$, 它具有性质

$$\dot{o}+\mathbf{x} \in S \Rightarrow \dot{o}+\lambda\mathbf{x} \in S, \quad \forall\lambda \in \mathbb{R}. \tag{6}$$

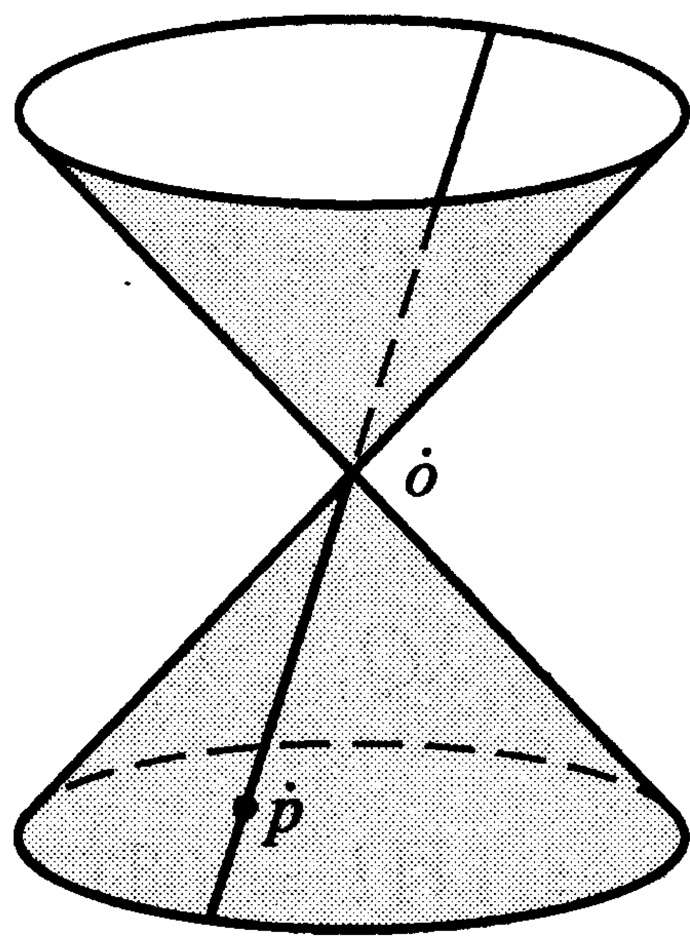


图11

在这种情形, 点$\dot{o}$可称为是**锥面的顶点**(它自动地成为对称中心), 而这些直线$\dot{o}+\lambda\mathbf{x}$就都被称之为**锥面的母线**. 只有$I'_{s,r}$型的二次曲面具有锥面性质(6)(在已给定的情形, 锥面的顶点就是坐标原点).

柱面S则可以作为这样的二次曲面(图12)加以刻画, 即有向量$\mathbf{u}\neq\mathbf{0}$, 使得

$$\dot{p}\in S\Rightarrow\dot{p}+\lambda\mathbf{u}\in S\qquad\forall\lambda\in\mathbb{R}.\tag{7}$$

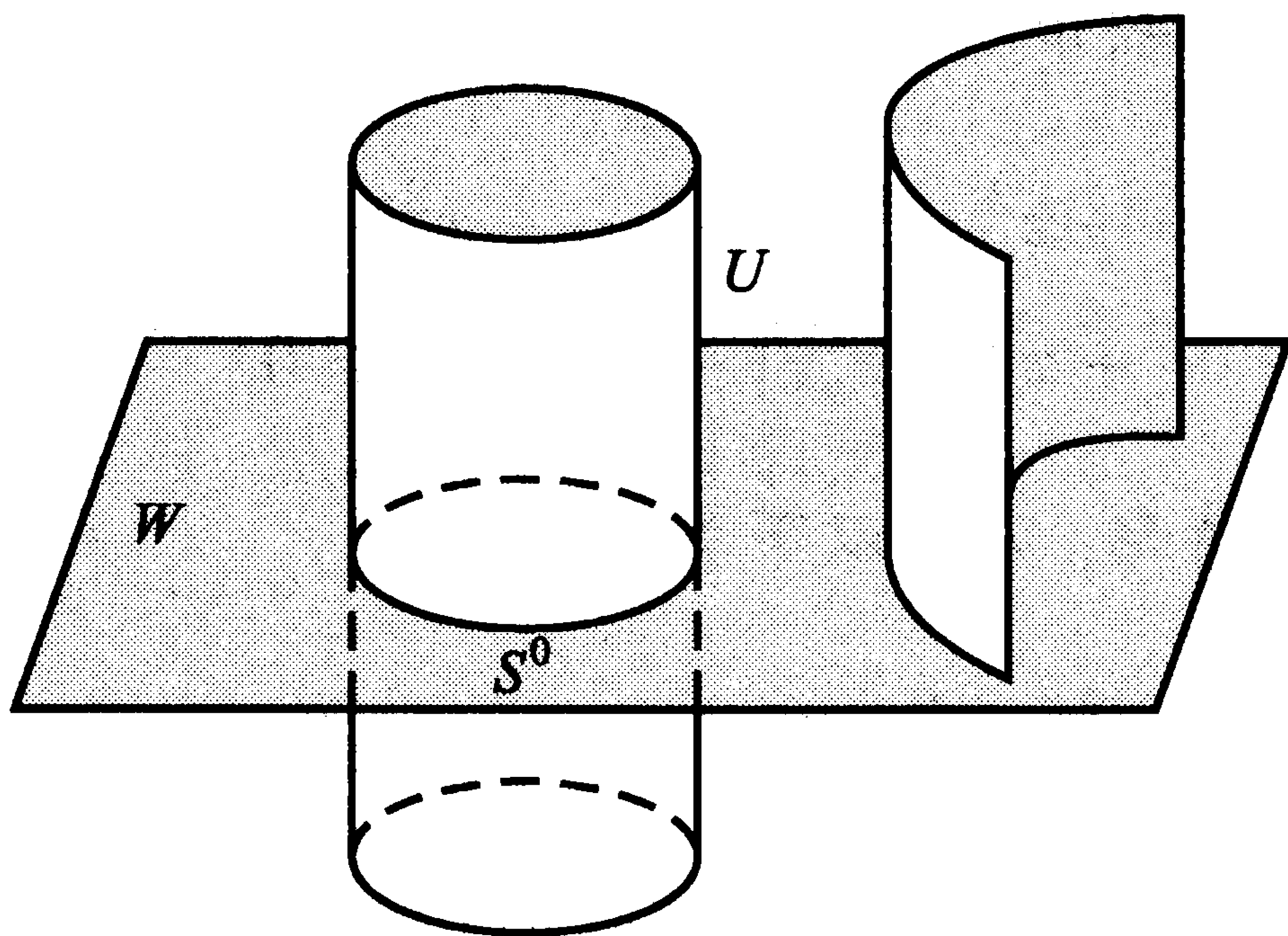


图12

换句话说, 沿着$\mathbf{u}$的平移$t_{\lambda\mathbf{u}}$把柱面S变成自己: $t_{\lambda\mathbf{u}}(S)=S$. 因为$t_{\mathbf{u}_1}\cdot t_{\mathbf{u}_2}=t_{\mathbf{u}_1+\mathbf{u}_2}$, 所以, 具有性质(7)的所有向量构成一个向量子空间$U\subset V$. 形如$\dot{p}+U$, $\dot{p}\in S$的平面被称为是**柱面S的母线**. 如果$V=U\oplus W$, $\dot{q}\in S$, 那么, 每个母线$\dot{p}+U$与平面$\dot{q}+W$相交于唯一一点$\dot{r}$ ($\overrightarrow{pq}=\mathbf{u}+\mathbf{w}$, $\mathbf{u}\in U$, $\mathbf{w}\in W$, 从而$\dot{p}+\mathbf{u}=\dot{q}-\mathbf{w}=\dot{r}$). 所以, 子空间$U\subset V$与柱面$S$的二次曲面

$$S^0=S\cap(\dot{q}+W)$$

是唯一确定的. 称二次曲面S^0是柱面S的**底面**.

如果$\dot{p}=\dot{o}+\mathbf{x}\in S_Q$且$\dot{p}+\alpha\mathbf{u}\in S_Q$, 也就是$Q(\dot{p})=0$且$Q(\dot{p}+\alpha\mathbf{u})=0$, 那么, 由§1的关系式(5), 我们有

$$q(\alpha\mathbf{u})+2\{f(\mathbf{x},\alpha\mathbf{u})+l(\alpha\mathbf{u})\}=0.$$

这就意味着

$$\alpha^2 q(\mathbf{u}) + 2\alpha\{f(\mathbf{x},\mathbf{u}) + l(\mathbf{u})\} = 0, \qquad \forall\alpha \in \mathbb{R},$$

从而有

$$q(\mathbf{u}) = 0, \qquad f(\mathbf{x},\mathbf{u}) + l(\mathbf{u}) = 0. \tag{8}$$

设$\mathbf{u} \in U$且

$$V = U \oplus W, \quad W = \langle \mathbf{e}_1, \cdots, \mathbf{e}_m, \rangle \quad U = \langle \mathbf{e}_{m+1}, \cdots, \mathbf{e}_n \rangle .$$

那么, 由关系式(8)可以得到, 在表达式

$$Q(\dot{o} + \mathbf{x}) = \sum_{i,j} x_i x_j + 2\sum_j \varphi_j x_j + \varphi_0$$

中$\varphi_{i,j}$和所有$j > m$的φ_j都等于零. 可见, 在二次曲面的标准方程中没有子空间U中基底向量对应的坐标. 这就得到如下的结论:

如果$r = \text{rank } Q$, 那么S_Q是个柱面$\Leftrightarrow$在可中心化的二次曲面情形$r < n$, 在不可中心化的情形$r < n - 1$. 其次, $\dim\ U = n - r$或相应地$n - r - 1$, 而柱面S_Q的底面S_Q^0是个非退化的二次曲面, 或者是个r维的仿射空间的锥面(可中心化的二次曲面), 或者是$r + 1$维的仿射空间锥面(不可中心化的二次曲面).

定义4 *根据柱面的底面的情况分别称之为***椭圆的, 双曲的***或***抛物的***. 同样, 可以讨论二次曲面S^0上的中心.*

应该注意到, 可以这样来区分锥面和柱面, 看它们的顶点是有限点还是无穷远点.

4. 二次曲面的类型 我们称$r = \text{rank}\, q$是二次函数Q以及其对应的二次曲面S_Q的秩, 其中q是与Q相关联的二次型. 人们常常称这个数是二次曲面S_Q的**小秩**, 还可以引进一个**大秩**与r并列在一起. 为了给出这个定义, 依照二次曲面在任意一个坐标系$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_n\}$之下的一般方程式

$$Q(\dot{o} + \mathbf{x}) = \sum_{i,j=1}^{n} \varphi_{ij} x_i x_j + 2\sum_{i=1}^{n} \varphi_i x_i + \varphi_0 = 0 \tag{9}$$

可以构造出两个矩阵, 即二次型q的矩阵$F = (\varphi_{ij})$和加边矩阵

$$\widetilde{F} = \begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1n} & \varphi_1 \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_{n1} & \cdots & \varphi_{nn} & \varphi_n \\ \varphi_1 & \cdots & \varphi_n & \varphi_0 \end{pmatrix}.$$

于是, 按定义, $r = \text{rank } F$, 而$\widetilde{r} = \text{rank } \widetilde{F}$. 为方便起见, 我们设

$$\varphi_{i,n+1} = \varphi_{n+1,i} := \varphi_i, \quad i = 1, 2, \cdots, n; \qquad \varphi_{n+1,n+1} := \varphi_0,$$

从而, $\tilde{F} = (\varphi_{ij})_1^{n+1}$且$\tilde{r}$就是二次型$\tilde{q}$:

$$\tilde{q}(\mathbf{x}) = \tilde{Q}(\dot{o} + \mathbf{x}) = \sum_{i,j=1}^{n+1} \varphi_{ij} x_i x_j$$

的秩.

可以认为$x_{n+1} = 1$. 坚持这些假定再由$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_n\}$向新的坐标架$\{\dot{o}';\ \mathbf{e}'_1, \cdots, \mathbf{e}'_n\}$做形式的转换, 我们得到

$$\begin{aligned} x_1 &= a_{11}x'_1 + \cdots + a_{1n}x'_n + a_{1,n+1}x'_{n+1}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= a_{n1}x'_1 + \cdots + a_{nn}x'_n + a_{n,n+1}x'_{n+1}, \\ x_{n+1} &= x'_{n+1} \end{aligned}$$

$(x_{n+1} = x'_{n+1} = 1)$. 转换矩阵$(a_{ij})_1^{n+1}$ 当然是非退化的. 显然, 在新的坐标之下, 二次型$\tilde{q}$的矩阵转化为

$$\tilde{F}' = {}^t\!A \cdot \tilde{F} \cdot A.$$

因为$\det\ A \neq 0$, 所以$\tilde{r}$是相对于仿射变换的一个不变量, 而且, 特别地, $\tilde{r}$很容易按着二次曲面S_Q在其标准形式下的方程式计算出来. 我们看到, 二次曲面S_Q是退化的, 也就是说, 如果$\tilde{r} < n+1$, 或者, 同样的, 如果$\det\ \tilde{F} = 0$, 则S_Q必为锥面或者柱面.

前面我们对于小秩r的仿射不变性已经加以关注. 标准方程式中以± 1为系数的平方项的个数由特征多项式$\chi_F(t)$的正根和负根的个数决定, 二次曲面的中心性可由条件$\det\ F \neq 0$推导出来. 如果$\det\ F = 0$, 那么, 或者没有中心($\operatorname{rank} F < \operatorname{rank} \widetilde{F}$), 或者有无穷多个($\operatorname{rank} F = \operatorname{rank} \widetilde{F}$), 后者对应柱面二次曲面的情形. 换言之, 二次曲面的标准形式可以完全刻画出来而不需要把它的方程式$Q(\dot{p}) = 0$地简化成标准形式.

我们再引进一些在研究二次曲面时以极其自然的方式产生的概念. 假如想找出二次曲面(9)与经过点$\dot{p}_0 = (x_1^0, \cdots, x_n^0)$的直线

$$x_i = x_i^0 + \alpha_i t, \qquad i = 1, 2, \cdots, n \tag{10}$$

的交集. 把(10)代入(9), 我们得到一个关于t的二次方程式

$$Q^{(0)}t^2 + 2Q^{(1)}t + Q^{(2)} = 0, \tag{11}$$

它的系数是

$$\begin{gathered} Q^{(0)} = q(\alpha), \\ Q^{(1)} = \sum_{i=1}^{n} Q_i(\dot{p}_0)\alpha_i, \qquad Q_i(\dot{p}) = \frac{1}{2}\frac{\partial Q}{\partial x_i} = \sum_{j=1}^{n} \varphi_{ij} x_j + \varphi_i, \\ Q^{(2)} = Q(\dot{p}_0). \end{gathered}$$

这里的$\alpha = (\alpha_1, \cdots, \alpha_n)$是直线(10)的一个方向向量, 而$x_1, \cdots, x_n$是流动(当前)点$\dot{p}$的坐标.

定义5 如果$q(\alpha) = 0$, 就称向量$\alpha = (\alpha_1, \cdots, \alpha_n)$是对于二次曲面$S_Q$的一个**渐近向量**. 方程式

$$q(\alpha) = 0$$

就给出了所谓的二次曲面S_Q的**渐近方向的锥面**.

如果直线(10)不是一个渐近方向, 即$q(\alpha) \neq 0$, 那么, 有两个根(可能是复共轭的), 它们对应直线和二次曲面的一对交点(可能是虚的). 渐近方向直线, 或者与二次曲面不相交, 或者相交于一点, 或者, 最后, 它完全包含在S_Q中(在最后一种情形, 直线(3)是二次曲面S_Q的直纹母线).

假设$\dot{p}_0 = (x_1^0, \cdots, x_n^0)$是二次曲面的一个点, 也就是$Q^{(2)} = Q(\dot{p}_0) = 0$. 如果对于$i = 1, \cdots, n$都有$Q_i(\dot{p}_0) = 0$, 则称$\dot{p}_0$是二次曲面$S_Q$的一个**奇点**. 奇点的坐标$x_1^0, \cdots, x_n^0$可由方程组

$$\sum_{j=1}^{n} \varphi_{ij} x_j + \varphi_i = 0, \qquad i = 1, 2, \cdots, n, \qquad \sum_{j=1}^{n} \varphi_j x_j + \varphi_0 = 0$$

得出来.

显然, 奇点只能在退化的二次曲面中出现, 而且, 当$\tilde{r} = n$时, 至多只能存在一个奇点. 在一般情形, 奇点在一个$n - \tilde{r}$维的平面上. 方程式

$$\sum_{i=1}^{n} Q_i(\dot{p}_0)(x_i - x_i^0) = 0$$

给出二次曲面S_Q在它的非奇异点的切平面.

5. 欧几里得空间中的二次曲面 设$\mathbb{E}$是个n维的欧几里得空间. V是与其相关联的$\mathbb{R}$上的向量空间. 如同在一般的仿射空间一样, 用方程$Q(\dot{p}) = 0$可给定二次曲面$S_Q \subset \mathbb{E}$.

§1中关于$\mathbb{E}$上二次函数的Iso($\mathbb{E}$)-等价性的定理3的一个显而易见的套用是

定理3 在n维欧几里得空间$\mathbb{E}$中, 二次曲面的方程式通过选择直角坐标系$\{\dot{o}; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$可以化成一个而且只能是一个标准型. 也就是说, 以坐标原点$\dot{o}$为对称中心的可中心化的二次曲面, 穷尽形式

$$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_s^2}{a_s^2} - \frac{x_{s+1}^2}{a_{s+1}^2} - \cdots - \frac{x_r^2}{a_r^2} = 1, \qquad 0 < s \leqslant r, \tag{12}$$

$$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_s^2}{a_s^2} - \frac{x_{s+1}^2}{a_{s+1}^2} - \cdots - \frac{x_r^2}{a_r^2} = 0, \qquad \frac{r}{2} \leqslant s < r; \tag{13}$$

而不可中心化的二次曲面穷尽形式

$$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_s^2}{a_s^2} - \frac{x_{s+1}^2}{a_{s+1}^2} - \cdots - \frac{x_r^2}{a_r^2} + 2x_{r+1} = 0, \qquad \frac{r}{2} \leqslant s \leqslant r \tag{14}$$

(零二次曲面和二重子空间除外).

应当把定理3的叙述加以补充, 描述一下量a_i. 在情形(12):

$$a_i = \sqrt{\left|\frac{\varphi_0}{\lambda_i}\right|} > 0, \qquad i = 1, \cdots, r, \tag{12'}$$

其中$0 \neq \varphi_0 = Q(\dot{o})$, 而$\lambda_1, \cdots, \lambda_r$是对称矩阵$F = (\varphi_{ij})$的特征根(固有值). 总可以选择变元$x_i$的标号顺序使得表达式

$$Q(\dot{o} + \mathbf{x}) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + \varphi_0$$

对于量λ_i, φ_0满足不等式

$$\lambda_1\varphi_0 < 0, \quad \cdots, \quad \lambda_s\varphi_0 < 0; \quad \lambda_i\varphi_0 > 0, \quad i > s.$$

在情形(13), 可以令

$$a_i = \frac{1}{\sqrt{|\lambda_i|}} > 0, \qquad i = 1, 2, \cdots, r, \tag{13'}$$

同时, 显然. 把必要性情形方程式的两端均乘以-1, 就总能满足条件$s \geqslant r/2$.

在情形(14), 我们可以认为

$$\lambda_1\mu > 0, \quad \cdots, \quad \lambda_s\mu > 0; \quad \lambda_i\mu < 0, \quad i > s,$$

所以,

$$a_i = \sqrt{\left|\frac{\mu}{\lambda_i}\right|} > 0, \quad i = 1, \cdots, r. \tag{14'}$$

非退化的二次曲面的仿射名称: **椭球面**((12), $s = n$, 图13), **双曲面**((12), $0 < s < r = n$, 图14, 图15), **椭圆抛物面**((14), $s = n - 1$, 图16), **双曲抛物面**((14), $0 \leqslant s < r = n - 1$, 图17)都可以照搬到欧几里得二次曲面上来, 可是, 出现了连续不变量(参变量)——所说的**半轴**a_i. 它们的Iso($\mathbb{E}$)-不变性是在§1的定理3中构建的量λ_i, φ_0, μ的Iso($\mathbb{E}$)-不变性的推论. 例如, 用仿射的观点看, 给定了n, 所有的椭球面都等价于单位球面. 从欧几里得空间的观点来看, 甚至球面也有自己的不变量, 就是半径$R = a_1 = \cdots = a_n$(具有相等的半轴的椭球). 以$a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n > 0$为半轴的椭球, 显然, 可以认为内切于以$a_1$为半径的球面, 因为中心到椭球上点$(x_1, \cdots, x_n)$的距离等于$\sqrt{x_1^2 + \cdots + x_n^2}$, 而且

$$1 = \frac{x_1^2}{a_1^2} + \cdots + \frac{x_n^2}{a_n^2} \geqslant \frac{1}{a_1^2}(x_1^2 + \cdots + x_n^2),$$

也就是$x_1^2 + \cdots + x_n^2 \leqslant a_1^2$, 而且在点$(a_1, 0, \cdots, 0)$处达到相等. 类似地, 以$a_n$为半径的球面内切于这个椭球面.

在双曲面情形, 半轴$a_{s+1}, \cdots, a_n$被称为**极小半轴**. 这个术语表达这样的事实, 平面$x_1 = \cdots = x_s = 0$的双曲面截口没有实点. 一般地, 借助截面研究二次曲面是通行的几何学方法, 更有利于在高维图形中给出具有直观性的元素.

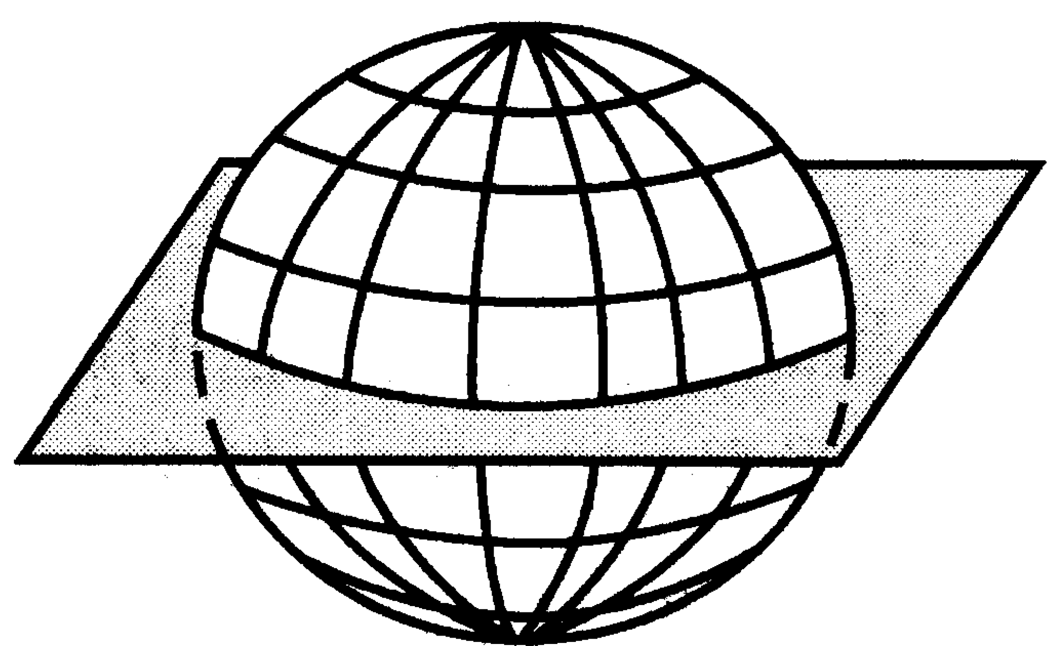


图13

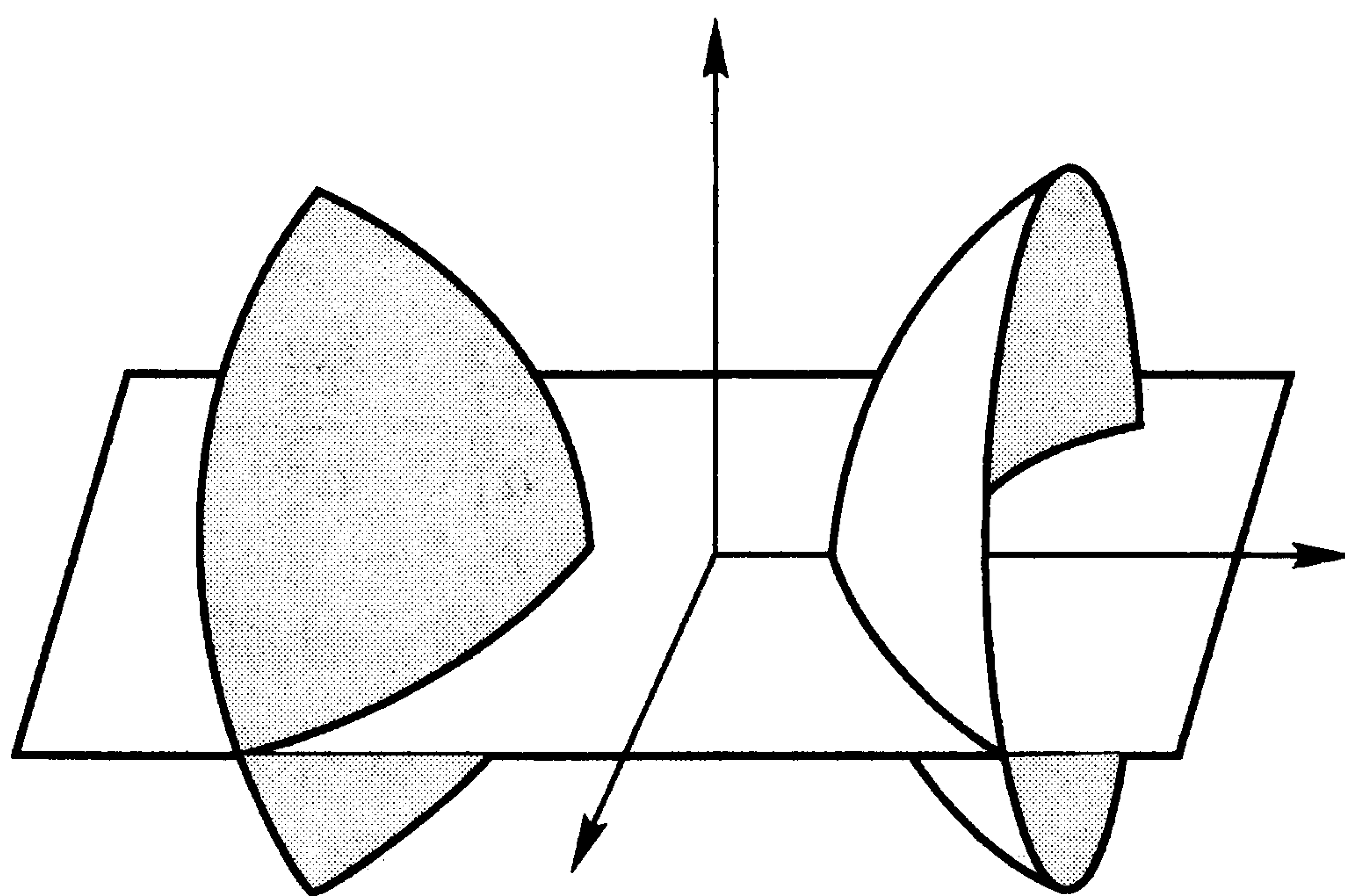


图14

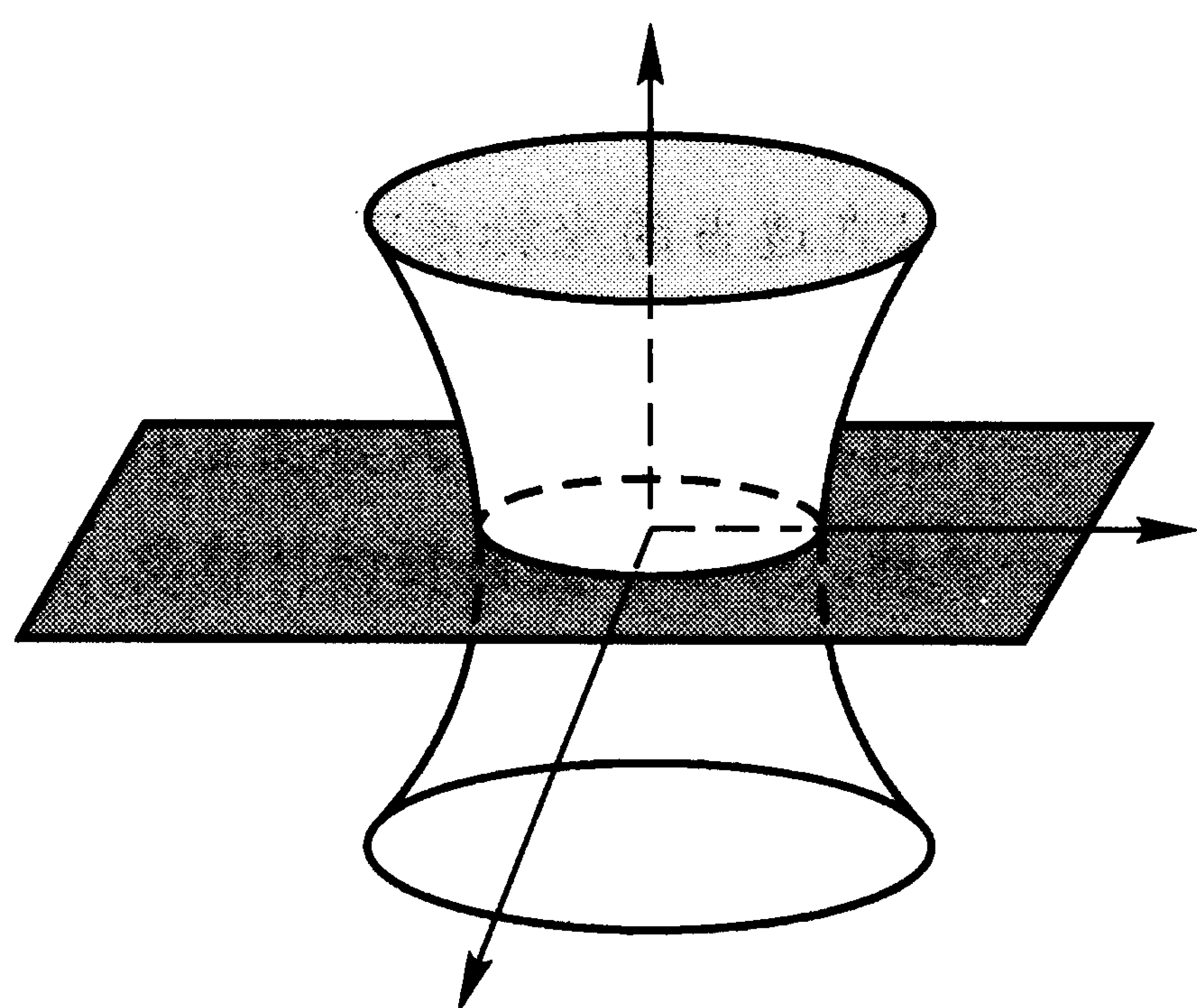


图15

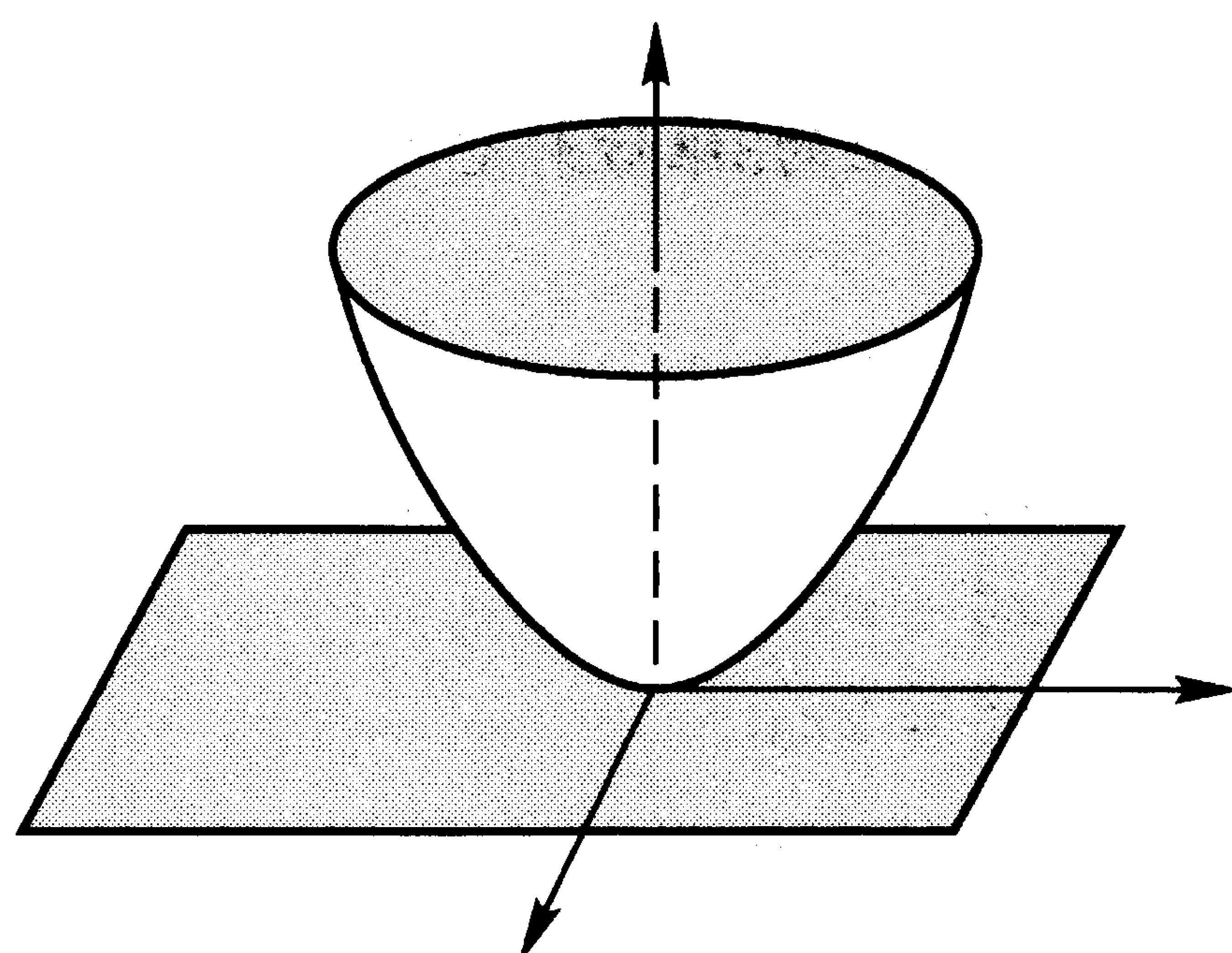


图16

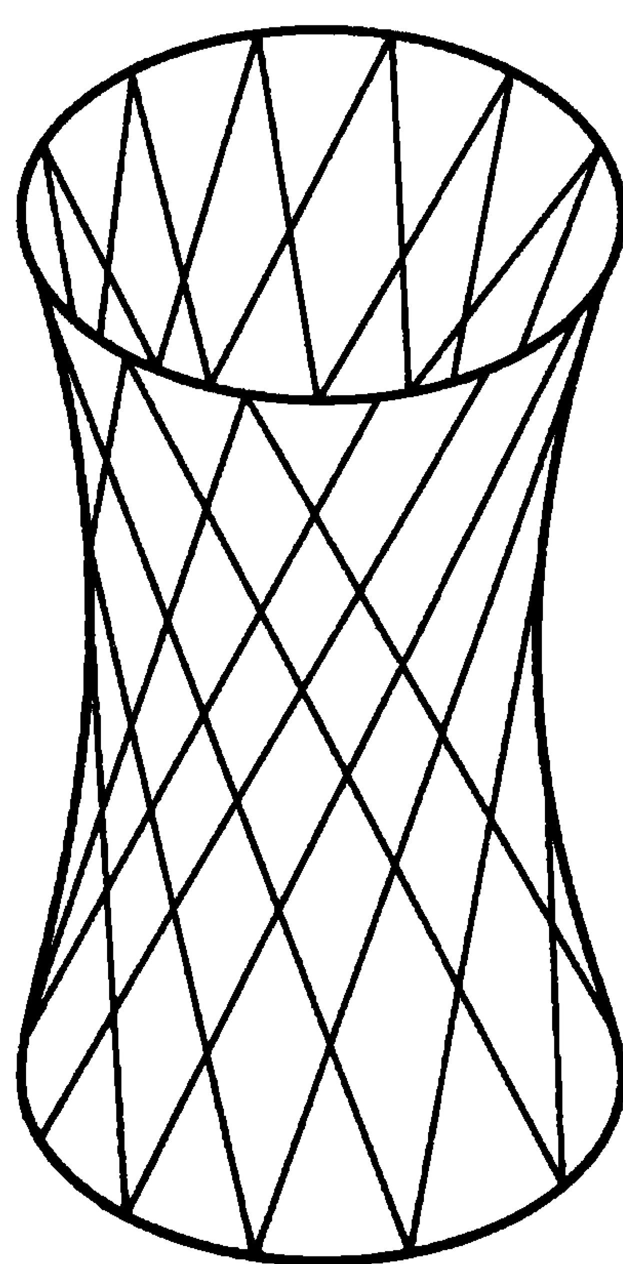


图17

用超平面x_i=常数$< a_i$截取椭球面, 就在$n-1$维空间中又给出了一个椭球面. 而双曲面的截面的差别可以说是五花八门的. 当$n=2$时, 双曲线$x_1^2/a_1^2 - x_2^2/a_2^2 = 1$, 当$n=3$时, 双叶双曲面和单叶双曲面

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1, \qquad \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1,$$

我们可以把它们表达在平面图上(图14, 图15).

当$n=4$的时候, 三种不同类型的双曲面借助截取可得到双曲面和椭球面. 比如, 在相对论中经常遇到的双叶双曲面

$$\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}-\frac{x_4^2}{a_4^2}=1,$$

就是由两个相联系的分支组成的, 分别在半空间$x_1 \geqslant a_1$和$x_1 \leqslant a_1$中. 它的超平面x_i=常数, $i>1$的截口通常是个双叶双曲面, 而超平面x_1=常数, $|x_1|>a_1$的截口给出一个椭球面. 有多种类型的抛物面, 我们既不停下来分析抛物面, 也不分析圆锥和圆柱(椭圆的, 双曲的和抛物的).

习 题

1. 证明, 在$\mathbb{R}$上的3维仿射空间$\mathbb{A}$中, 任意二次曲面都可以在适当的坐标系下, 由下列的方程式之一给出来:

(1) $x_1^2+x_2^2+x_3^2=1$;

(2) $x_1^2+x_2^2-x_3^2=1$;

(3) $x_1^2+x_2^2-x_3^2=-1$;

(4) $x_1^2-x_2^2=2x_3$;

(5) $x_1^2+x_2^2=2x_3$;

(6) $x_1^2+x_2^2+x_3^2=-1$;

(7) $x_1^2+x_2^2-x_3^2=0$;

(8) $x_1^2+x_2^2+x_3^2=0$;

(9) $x_1^2+x_2^2=-1$;

(10) $x_1^2+x_2^2=1$;

(11) $x_1^2=2x_2$;

(12) $x_1^2-x_2^2=1$;

(13) $x_1^2-x_2^2=0$;

(14) $x_1^2-1=0$;

(15) $x_1^2+x_2^2=0$;

(16) $x_1^2+1=0$;

(17) $x_1^2=0$.

2. 设S是欧几里得点空间$\mathbb{E}$的一个二次曲面. 如果在直角坐标架$\{\dot{o};\ \mathbf{e}_1, \cdots, \mathbf{e}_n\}$之下它的方程式是

$$\sum_{i=1}^{n}\lambda_i x_i^2+2\sum_{i=1}^{n}\mu_i x_i+\mu_0=0,$$

那么, 就可以讨论二次曲面S的**主方向标架**(回忆一下二次型的“主轴”). 试找出二次曲面

(1) $2x^2+y^2-3z^2+12xy+4xz+8yz+18=0$;

(2) $6x^2+5y^2+7z^2+4xy-4xz-8x-10y+14z-6=0$

的主方向.

3. 实二次型$q(\mathbf{x})$极值问题中，习题3.3.8可以给出本质的几何意义. 我们回忆一下，当$\|\mathbf{v}\|=1$时，值$q(\mathbf{v})$的意义，或者$q(\mathbf{v})/\|\mathbf{v}\|^2$的意义. 就可以验证

$$\max_{\|\mathbf{v}\|=1} q(\mathbf{v}) = \min_{q(\mathbf{v})} \|\mathbf{v}\|^2$$

(当交换min和max位置时，可得类似的结果). 比方说，设$n=2$且$q(\mathbf{v})=\alpha x^2+2\beta xy+\gamma y^2=1$是个椭圆方程式. 如果$\lambda_i$, $i=1, 2$是特征多项式$\lambda^2-(\alpha+\gamma)\lambda+(\alpha\gamma-\beta^2)=0$的根，且$(x_i, y_i)$对应实的极值点. 那么，$x_i^2+y_i^2=1/\lambda_i$，也就是说，其中一个根对应坐标原点到椭圆的最小距离的平方(第二个根对应最大距离的平方). 同时，可以得到椭圆主轴(主方向)的正交性以及关于椭圆面积的公式$\pi\Big/\sqrt{\alpha\gamma-\beta^2}$ (说明理由!).

4. 变量t取怎样的值时，二次曲面

$$x_1^2+x_2^2+x_3^2+2tx_1x_2+2tx_1x_3+2tx_2x_3-4t=0$$

是个椭圆面.

5. 找出二次曲面

$$x_1^2+5x_2^2+x_3^2+2x_1x_2+2x_2x_3+6x_1x_3-2x_1+6x_2+2x_3=0$$

和平面$2x_1-x_2+x_3=0$相交的曲线的仿射类型.

6. 什么时候两个双曲面有共同的渐近顶点.

7. 莫斯科的舒赫夫塔楼令人想起何种二次曲面?

§3 射影空间

射影几何的发展，特别是在19世纪的上半叶，对整个数学都产生了本质上的影响. 我们将只涉及与它有关的不太多的事实，可以参阅教学参考书[2]或专门的文献加以补充.

1. 射影平面的模型 在域$\mathcal{K}$的仿射平面上，任意两点均可位于唯一的一条直线上，而任意两条不平行的直线必相交于一点.

让我们回想一下在解析几何中射影平面$\mathbb{P}^2=\mathcal{K}\mathbb{P}^2$的构造，在它上面

i) 任意两个不同的点必然位于唯一的一条直线上;

ii) 任意两条不相同的直线必相交于一点.

为了构造$\mathbb{P}^2$，我们从域$\mathcal{K}$上任意一个3维向量空间V开始，且定义$\mathbb{P}^2=\mathbb{P}(V)$，认为点$p\in\mathbb{P}(V)$是$V$的1维子空间(直线)，而直线$L\subset\mathbb{P}(V)$是$V$的2维子空间. 如果仿射子空间$p$被包含在$L$中，那么，点$p$就位于射影直线$L$上(或者与$L$关联). 关联性i)显然是满足的: 如果$p\neq q$都是点，那么，它们在$V$中就是不同的直线，它们的和必然是个2维子空间$L$，也就是$\mathbb{P}(V)$的一条直线. 其次，两条不同的射影直线$L$和$M$本来在$V$中就是不同的2维子空间，所以它们在$V$中的和$L+M$就应该是整个空间$V$. 进而，由第一章§2的公式(7)，我们有

$$\dim(L\cap M)=\dim L+\dim M-\dim(L+M)=2+2-3=1.$$

这意味着，$L\cap M$是个1维的仿射子空间，也就是有唯一的一点$p\in\mathbb{P}(V)$，射影直线L和M相交于p. 关联性ii)可以用同样的办法被满足.

根据上面的叙述, 可以导出射影平面的一种实现方式, 为了接近它, 我们索性将其称为仿射空间的**束**. 我们用下面的模型来获得射影平面刻画的已知的直观性.

设$\mathfrak{K}=\mathbb{R}$是实数域. 在欧几里得空间$\mathbb{E}=\mathbb{R}^3$中选取2维球面

$$S^2: x^2+y^2+z^2=1.$$

经过$\mathbb{R}^3$的原点$\dot{o}$的每条直线和单位球面两个径对点, 而每个包含原点$\dot{o}$的平面与球面相交都得一个大圆.

在射影空间$\mathbb{RP}^2$中, 可以选择径对点对(t,t')为点(图18), 而选择S^2上的大圆为直线, 当大圆周与直线L相交于t和t'时, 就认为$p=(t,t')$属于L. 完全是显然的, 两个不同的大圆周只能相交于一个径对点对(t,t'). 关联性i)和ii)都得到了满足.

可以把我们的研究限制在下半球面S^2_-上, 它由所有满足$x^2+y^2+z^2=1$, $z\leqslant 0$的点(x, y, z)组成(图19). 它的界面就是满足方程式$x^2+y^2=1$且$z=0$的赤道$S^1\subset S^2$. 球面S^2的径对点对中至少有一个点属于S^2_-, 而且只有它们是赤道S^1的径对点才能使点对的两个点同时属于S^2. 换言之, 点$p\in\mathbb{RP}^2$是本球面S^2_-的一个点, 前提条件是把赤道上的径对点对当成一个点. 在$\mathbb{RP}^2$上认为直线L就是S^2上的大圆周与S^2_-的交集. 特别地, 赤道S^1在把径对点等同起来之后本身就是一条直线.

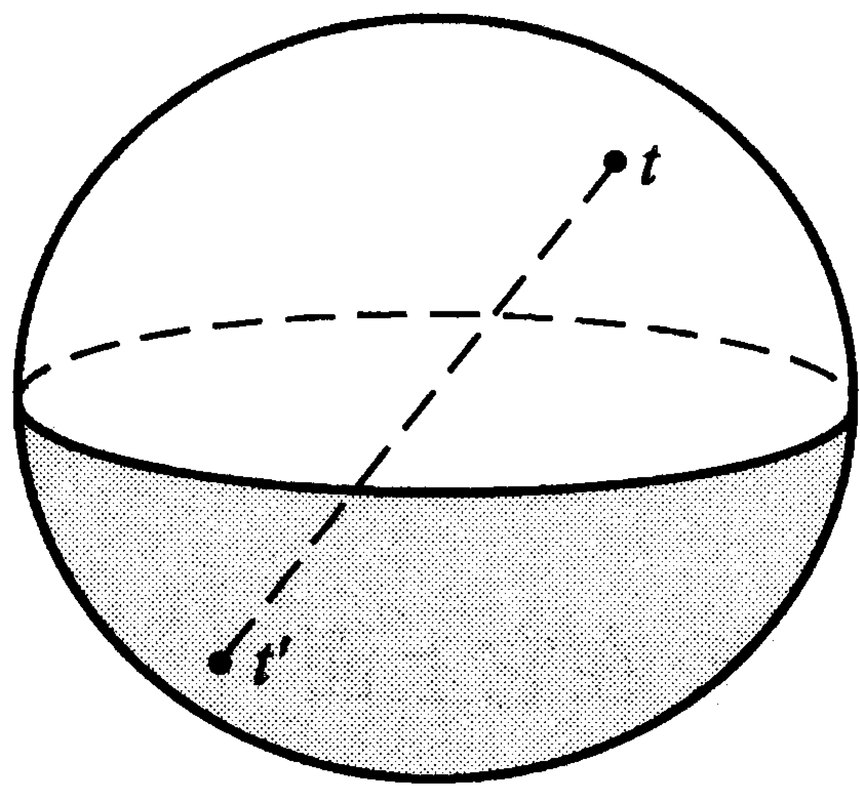


图18

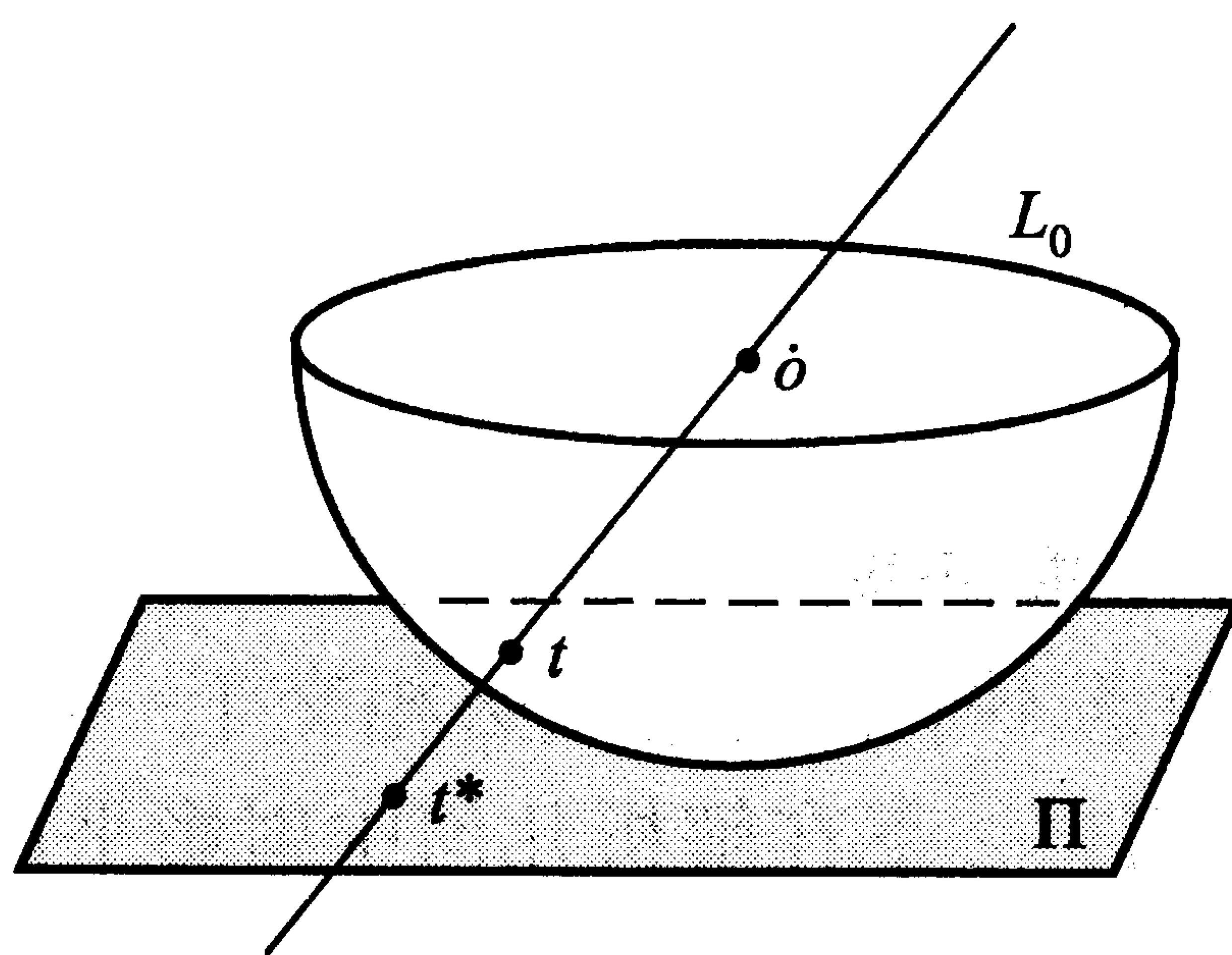


图19

考察仿射平面Π, 它与半球面S^2_-相切于南极点$(0, 0, -1)$, 这个以$\dot{o}$为原点的半球面树立在Π上. 这意味着, 点$t\in S^2_-\backslash S^1$与位于连接点$\dot{o}$和点t的直线上的点$t^*\in\Pi$相对应. 显然, 射影

$$\sigma: S^2_-\backslash S^1\to\Pi$$

是双向单值的. 映射σ把$\mathbb{RP}^2$上的每一条直线, 也就是S^2上的大圆周的弧, 对应Π上的直线. 映射σ^{-1}把点变成点时保持关联性, 映射σ^{-1}的像就是$\mathbb{RP}^2$中除了代表半球S^2的赤道S^1的直线L_0上的点以外的所有的点的集合. 这样一来, 射影平面可以由仿射平面得到, 只要补充某些新的投影直线L_0, 即所说的**无穷远直线**. 仿射平面上任意平行直线的集合必然是S^2_-经过赤道的直径的终点的大圆周弧的集合在σ^{-1}之下映过来的. 于是, 赤道上径对应的点对(t,t'), 按约定, 也就确定了是直线L_0. 可见, 在仿射平面

上补充的无穷远直线也就是在Π上补充了在它上面不相交的平行直线对的交点. 必须注意, 无穷远直线L_0不是仿射直线.

与射影平面不同, 射影直线$\mathbb{RP}^1$是经过普通平面固定点$\dot{o}$的直线束(图20). 在固定的经过点$\dot{o}$的圆周S^1上, 每个直线束对应一个点(它和S^1的交点). 点$\dot{o}$对应S^1在$\dot{o}$点的切线. 也就是说, 圆周就是射影直线的一个模型.

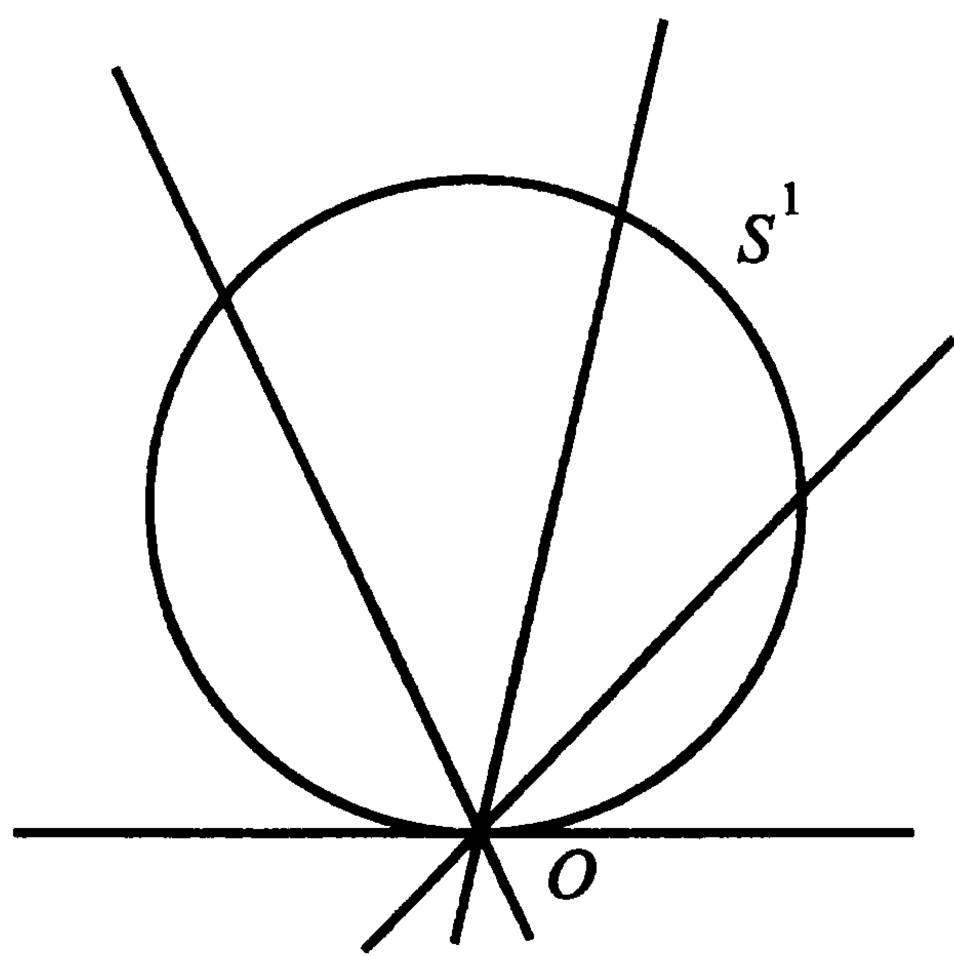


图20

2. 任意维的射影空间 完成了射影直线和射影平面的直观表达, 就可以很容易地引入在任意域上的维数更高的射影空间的概念. 可以把它理解为一个点集, 它带有某些可分离的子集(可称为**射影子空间**), 它们服从一些自然的公理或者关联性关系. 公理化方法(也称**综合法**)有自己的优越性, 但太过于迂回了, 而且也不太适应其他教程的内容. 因此, 我们选择一个直接的入口, 本质上等价于与仿射空间连起来加以研究.

定义1 把域$\mathfrak{K}$上$n+1$维向量空间V的(齐次的或向量的)直线的集合$\mathbb{P}^n = \mathfrak{K}\mathbb{P}^n = \mathbb{P}(V)$称为域$\mathfrak{K}$上的$n$维**射影空间**. 把空间$V$的直线称为空间$\mathfrak{K}\mathbb{P}^n$的点. 如果$U \subset V$是$V$的一个$m+1$维的子空间, 那么, V中所有包含在U中的直线组合的集合$\mathbb{P}(U) \subset \mathbb{P}(V)$被称之为空间$\mathbb{P}^n$的$m$维**射影子空间**(也说是**射影线性流形**或**射影平面**). 当$m = n-1$时, 就称为**射影超平面**. 可以认为, $\mathbb{P}(\{\mathbf{0}\}) = \varnothing$是空集.

也可以用另外的方式表达这同一个定义, 例如说, 在V的对于$\{\mathbf{0}\}$的余集V^*上依据等价关系

$$\mathbf{x} \sim \mathbf{y} \Leftrightarrow \exists \lambda \in \mathfrak{K}^*, \quad \mathbf{x} = \lambda \mathbf{y},$$

做出商集$\mathbb{P}(V)$, 即可称之为$\mathfrak{K}$上向量空间V的向量生成的**射影空间**.

由$\mathbf{x} \in V^*$($\mathbf{x} \in V$是个非零向量)决定的等价类$\tilde{\mathbf{x}}$就是射影空间$\mathbb{P}(V)$的一个点. 换句话说, 按定义

$$\widetilde{\lambda \mathbf{x}} = \tilde{\mathbf{x}}, \quad \forall \lambda \in \mathfrak{K}^*. \tag{1}$$

映射$\Pi : \mathbf{x} \mapsto \tilde{\mathbf{x}}$被称之为$V^*$到它的商集$\mathbb{P}(V)$上的**标准映射**. 应当强调, 在$\mathbb{P}(V)$上并没有定义线性运算, 例如, 我们就不能规定$\tilde{\mathbf{x}} \oplus \tilde{\mathbf{y}} = \widetilde{\mathbf{x}+\mathbf{y}}$. 2维向量子空间$U \subset V$可以定义$\mathbb{P}(U)$的射影直线, 而3维向量子空间就定义的是射影平面. 如果$U \subset W$, 即U是

另外一个子空间$W \subset V$的向量子空间, 那么, $\mathbb{P}(U) \subset \mathbb{P}(W)$, 因为$V$中的每一条被包含在$U$中的直线一定也被包含在$W$中. 如果$\mathbb{P}(U) \subset \mathbb{P}(W)$, 则说射影子空间$\mathbb{P}(U)$属于$\mathbb{P}(W)$或者与$\mathbb{P}(W)$**关联**. 令

$$\mathbb{P}(U) \cap \mathbb{P}(U') = \mathbb{P}(U \cap U')$$

是有意义的.

对于每个集合$S \subset \mathbb{P}(V)$必然存在一个最小的包含S的射影子空间: 如果$S = \{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \cdots\}$, 那么$U = \langle \mathbf{x}_1, \mathbf{x}_2, \cdots \rangle_{\mathfrak{K}}$. 通常称$S$是$\mathbb{P}(U)$的一个**生成系**. 包含$\mathbb{P}(U)$和$\mathbb{P}(U')$的最小的射影子空间, 显然应该是$\mathbb{P}(U + U')$. 每个子空间$\mathbb{P}(W)$均可看成一个独立的射影空间, 对于$U \subset W$, 可派生出子集$\mathbb{P}(U)$.

3. 齐次坐标　设$(\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_n)$是向量空间$V$的一个基底. 如果

$$\mathbf{x} = \xi_0 \mathbf{e}_0 + \xi_1 \mathbf{e}_1 + \cdots + \xi_n \mathbf{e}_n \in V^*,$$

那么, 我们直接把$\xi_0, \xi_1, \cdots, \xi_n$称为点$\tilde{\mathbf{x}}$相对于基底$(\mathbf{e}_i)$的**齐次坐标**. 域$\mathfrak{K}$中任意$n+1$个不同时为零的元素组$(\xi_i)$必定是$\mathbb{P}(V)$中某个点相对于基底$(\mathbf{e}_i)$的齐次坐标组. 两个这样的组$(\xi_i)$, (μ_i)都是$\mathbb{P}(V)$中同一个点相对于同一个基底$(\mathbf{e}_i)$的齐次坐标组, 当且仅当, 在$\mathfrak{K}$中有某个$\lambda \neq 0$, 使得$\mu_i = \lambda \xi_i$, $i = 1, 2, \cdots, n$. 这个事实, 我们将用记法

$$\tilde{\mathbf{x}} = (\xi_0 : \xi_1 : \cdots : \xi_n)$$

来表达, 它表明, 在给定的基底之下, $\tilde{\mathbf{x}} \in \mathbb{P}(V)$与相互平行的齐次坐标组等价类之间有相互单值的对应.

如果$(\mathbf{e}'_0, \mathbf{e}'_1, \cdots, \mathbf{e}'_n)$是$V$的任意一个另外的基底, 同时,

$$\mathbf{e}'_j = \sum_{i=0}^{n} a_{ij} \mathbf{e}_i, \quad 0 \leqslant j \leqslant n,$$

那么, $(\xi'_0, \xi'_1, \cdots, \xi'_n)$是$\tilde{\mathbf{x}}$对于基底$(\mathbf{e}'_i)$的齐次坐标组, 当且仅当, 能找到$\lambda \in \mathfrak{K}^*$, 使得

$$\lambda \xi_i = \sum_{i=0}^{n} a_{ij} \xi'_j, \quad 0 \leqslant i \leqslant n. \tag{2}$$

实际上,

$$(\xi_0 : \xi_1 : \cdots : \xi_n) = \tilde{\mathbf{x}} = (\xi'_0 : \xi'_1 : \cdots : \xi'_n),$$

这只要回顾一下第1章由新坐标向原坐标的转换法则就足够了.

我们还应注意, 由定理4(第1章§3)得到的, 在给定的基底之下, 所有的子空间$\mathbb{P}(U) \subset \mathbb{P}(V)$都可以由齐次线性方程组

$$\begin{aligned} a_{10}\xi_0 + a_{11}\xi_1 + \cdots + a_{1n}\xi_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{r0}\xi_0 + a_{r1}\xi_1 + \cdots + a_{rn}\xi_n &= 0 \end{aligned} \tag{3}$$

给出.

4. 仿射图 在以$(\mathbf{e}_i)$为基底的向量空间V中, 我们选出向量子空间

$$V_0 = \langle \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n \rangle,$$

而在仿射空间$(\mathbb{E} = V, V)$中选出超平面

$$\mathbb{E}_0 = \mathbf{e}_0 + V_0 = \{\mathbf{e}_0 + \mathbf{x} \,|\mathbf{x} \in V_0\}.$$

正如我们已经知道的, 如果, 对$\dot{a} = \mathbf{a} = \mathbf{e}_0 + \mathbf{a}'$, $\dot{b} = \mathbf{e}_0 + \mathbf{b}'$, 令

$$\overrightarrow{ab} = \mathbf{b}' - \mathbf{a}',$$

则对$(\mathbb{E}_0, V_0)$就是个仿射空间. 不包含在V_0中的直线$\langle \mathbf{x} \rangle \subset V$与$\mathbb{E}_0$相交于唯一一点. 实际上,

$$\mathbf{x} \notin V_0 \Rightarrow \mathbf{x} = \xi_0 \mathbf{e}_0 + \cdots + \xi_n \mathbf{e}_n, \quad \xi_0 \neq 0.$$

这就意味着$\lambda \mathbf{x} = \lambda \xi_0 \mathbf{e}_0 + \cdots + \lambda \xi_n \mathbf{e}_n = \mathbf{e}_0 + \mathbf{y}$, $\mathbf{y} \in V_0$的充要条件是$\lambda \xi_0 = 1$.

建立直线$\langle \mathbf{x} \rangle$到这个交点$\langle \mathbf{x} \rangle \cap \mathbb{E}_0$的对应. 我们得到一个直线$\langle \mathbf{x} \rangle \not\subset V_0$与仿射空间$\mathbb{E}_0$之间的一个双射对应

$$\Phi : \langle \mathbf{x} \rangle \mapsto \langle \mathbf{x} \rangle \cap \mathbb{E}_0.$$

换句话说, Φ引出双射映射

$$\Phi_0 : \mathbb{P}(V) \setminus \mathbb{P}(V_0) \to \mathbb{E}_0. \tag{4}$$

把$\mathbb{P}(V) \setminus \mathbb{P}(V_0)$理解为以$\mathbb{P}(V_0)$为远离(抛出)超平面的射影平面$\mathbb{P}(V)$.

定义2 把仿射空间$\mathbb{E}_0$与映射Φ_0一起(但, 有时就简明地说是$\mathbb{P}(V) \setminus \mathbb{P}(V_0)$)和$\mathbb{E}_0$等同起来, 称之为射影空间$\mathbb{P}(V)$的仿射图. 这时, 就把$\mathbb{P}(V_0)$称为相对于图$\mathbb{E}_0$的**无穷远超平面**. $\mathbb{P}(V_0)$所包含的点和平面同样都被认为是无穷远的.

我们选择坐标. 按照$\mathbb{P}(V) \setminus \mathbb{P}(V_0)$的本意, 它由$\tilde{\mathbf{x}} = (\xi_0 : \xi_1 : \cdots : \xi_n)$且$\xi_0 \neq 0$组成. 在$\mathbb{E}_0$中选择仿射坐标系$\{\dot{e}_0; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$, 这里$\dot{e}_0 = \mathbf{e}_0$理解为$\mathbb{E}_0$中的点, 而把$\mathbf{e}_1, \cdots, \mathbf{e}_n$理解为和$\mathbb{E}_0$相关联的向量空间$V_0$的一个基底. 为了找出$\tilde{\mathbf{x}}$的仿射坐标, 需要找出直线$\langle \mathbf{x} \rangle = \langle \xi_0 \mathbf{e}_0 + \xi_1 \mathbf{e}_1 + \cdots + \xi_n \mathbf{e}_n \rangle$与$\mathbb{E}_0$的交点. 我们已经看到, 这个点形如

$$\mathbf{e}_0 + \frac{\xi_1}{\xi_0} \mathbf{e}_1 + \cdots + \frac{\xi_n}{\xi_0} \mathbf{e}_n.$$

这意味着, 在图$\mathbb{E}_0$的坐标系$\{\dot{e}_0; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$之下, 点$\tilde{\mathbf{x}}$的仿射坐标就是$\xi_1/\xi_0, \cdots, \xi_n/\xi_0$.

这样一来, 把$\tilde{\mathbf{x}} \in \mathbb{P}(V) \setminus \mathbb{P}(V_0)$看成点$\Phi_0(\tilde{\mathbf{x}}) \in \mathbb{E}_0$的坐标, 我们就得到$\mathbb{P}(V)$的仿射(非齐次)坐标系, 的的确确, 它只定义在$\mathbb{P}(V) \setminus \mathbb{P}(V_0)$上. 在这个集合的点和它们的非齐次坐标之间存在一个双射对应.

如果, U是V的一个$m+1$维的向量子空间, 那么, m维的射影平面$\mathbb{P}(U)$或者相对于$\mathbb{E}_0$是无穷远离的($U \subset V_0$)的情形, 或者它的像

$$\Phi_0(\mathbb{P}(U)) = U \cap \mathbb{E}_0 = \mathbf{e}_0 + U_0$$

是图$\mathbb{E}_0$上的m维仿射平面. 另一方面, 任意一个m维仿射平面$\mathbf{e}_0 + U_0 \subset \mathbb{E}_0$对应$m$维射影平面$\mathbb{P}(U)$, 其中$U = \langle \mathbf{e}_0 U_0 \rangle$. 推导出来的这个论断表明, Φ_0不仅仅是$\mathbb{P}(V) \setminus \mathbb{P}(V_0)$与$\mathbb{E}_0$之间的一个双射的点的对应, 而且也是同一维数平面之间的一个对应. 在这个意义之下, $\mathbb{P}(V)$可以由$\mathbb{E}_0$补充无穷远离的超平面得到.

用向量$\mathbf{e}_i$替换$\mathbf{e}_0$, 而用超平面

$$V_i = \langle \mathbf{e}_0, \cdots, \mathbf{e}_{i-1},\ \mathbf{e}_{i+1},\ \cdots, \mathbf{e}_n \rangle$$

替换V_0, 我们就得到另外的仿射图$(\mathbb{E}_i, \Phi_i)$. 它要求点$(\xi_0 : \xi_1 : \cdots : \xi_n)$中$\xi_i \neq 0$. 在以$\{\dot{e}_i; \mathbf{e}_0, \cdots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \cdots, \mathbf{e}_n\}$为坐标系的$\mathbb{E}_i$中, 点$(\xi_0 : \xi_1 : \cdots : \xi_n)$的仿射坐标就是

$$\left(\frac{\xi_0}{\xi_i}, \cdots, \frac{\xi_{i-1}}{\xi_i}, \frac{\xi_{i+1}}{\xi_i}, \cdots, \frac{\xi_n}{\xi_i}\right).$$

取遍序列$\mathbf{e}_0, \mathbf{e}_1, \cdots,$ 中其余的向量, 我们就得到$n+1$个图

$$(\mathbb{E}_i, \Phi_i), \quad i = 0, 1, \cdots, n.$$

它们的并集与$\mathbb{P}(V)$"重合". 事实上, 对任意点$\tilde{\mathbf{x}} = (\xi_0, \xi_1, \cdots, \xi_n) \in \mathbb{P}(V)$, 它至少有一个坐标$\xi_i$不等于零, 这就意味着$\Phi_i(\tilde{\mathbf{x}}) \in \mathbb{E}_i$. 按照把$\mathbb{P}(V) \setminus \mathbb{P}(V_i)$与$\mathbb{E}_i$等同起来的作法, 我们有

$$\mathbb{P}^n = \mathbb{P}(V) = \bigcup_{i=0}^{n} \mathbb{E}_i.$$

容易看出, 用个数更少的图是不能将$\mathbb{P}^n$覆盖的.

5. 代数(流形)簇的概念 我们说, 多项式

$$f(t_0, t_1, \cdots, t_n) \in \mathfrak{K}[t_0, t_1, \cdots, t_n]$$

在点$\tilde{\mathbf{x}} = (\xi_0, \xi_1, \cdots, \xi_n) \in \mathbb{P}(V)$处转化为零. 如果, $f(\xi_0, \xi_1, \cdots, \xi_n) = 0$. 这意味着, 对所有$\lambda \in \mathfrak{K}, \lambda \neq 0$都有$f(\lambda\xi_0, \lambda\xi_1, \cdots, \lambda\xi_n) = 0$. 设

$$f = f_0 + f_1 + \cdots + f_m,$$

其中f_i是f中所有i次单项式之和, 我们可以看出, 由条件

$$\begin{aligned} 0 &= f(\lambda\xi_0, \cdots, \lambda\xi_n) \\ &= f(\xi_0, \cdots, \xi_n) + \lambda f_1(\xi_0, \cdots, \xi_n) + \cdots + \lambda^m f_m(\xi_0, \cdots, \xi_n) \end{aligned}$$

在域$\mathfrak{K}$为无穷域的情形, 必然得到$f_i(\xi_0, \xi_1, \cdots, \xi_n) = 0$对任意$i = 0, 1, \cdots, m$都成立. 由此可见, 如果$f$在某一个点$\tilde{\mathbf{x}} \in \mathbb{P}(V)$转化为零, 那么, 组成$f$的所有的单项式在同一点也都转化为零. 因此, 自然地引进下面的

定义3 满足代数方程组

$$g_1(\alpha_0, \cdots, \alpha_n) = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$g_k(\alpha_0, \cdots, \alpha_n) = 0,$$

的点$(\alpha_0 : \cdots : \alpha_n)$的子集合$S \subset \mathbb{P}^n$被称为(射影的)**代数流形**(亦称**代数簇**). 其中$g_1, \cdots, g_k$是齐次多项式.

更准确地, 可以研究$\mathbb{P}^n$中的闭的代数子集, 因为流形可适当地在齐次单纯理想扎里斯基拓扑语言基础上引入. 但是, 我们不去细说它. 代数(流形)簇(特别地, 当$\mathfrak{K} = \mathbb{C}$时, 复的代数(流形)簇)是一个更大的独立的数学分支的研究对象, 就是代数几何学.

我们只限于一个方程式

$$g(\alpha_0, \alpha_1, \cdots, \alpha_n) = 0.$$

的简单情形. 找出交集$S_0 = S \cap \mathbb{E}_0$的方程式. 如果$\dot{x} \in \mathbb{E}_0$, $\tilde{\mathbf{x}} = (\alpha_0 : \alpha_1 : \cdots : \alpha_n)$, 那么, $\alpha_0 \neq 0$. 因此, 条件$g(\alpha_0, \cdots, \alpha_n) = 0$等价于条件

$$g\left(1, \frac{\alpha_1}{\alpha_0}, \cdots, \frac{\alpha_n}{\alpha_0}\right) = 0.$$

因为$\alpha_1/\alpha_0, \cdots, \alpha_n/\alpha_0$是点$\dot{x}$的仿射坐标, 所以, 这就是"仿射"流形$S_0$的方程式, 也就是流形$S$在$\mathbb{E}_0$的方程式. 类似地, 在$\alpha_i$上实施, 就得到$S$在$\mathbb{E}_i$中的方程式.

反过来: 如果用$x_1, \cdots, x_n$代表$\mathbb{E}_0$中点的坐标且由方程式

$$f(x_1, \cdots, x_n) = 0$$

给出S_0, 其中f是任意的, 不一定是m次齐次多项式, 那么

$$g(\alpha_0, \alpha_1, \cdots, \alpha_n) = (\alpha_0)^m f\left(\frac{\alpha_1}{\alpha_0}, \cdots, \frac{\alpha_n}{\alpha_0}\right)$$

就是个齐次多项式. 实际上, 由f中的单项式

$$(x_1)^{k_1} \cdots (x_n)^{k_n}, \quad k_1 + \cdots + k_n \leqslant m,$$

就可以得到g中的m次单项式$(\alpha_0)^{m-k_1-\cdots-k_n}(\alpha_1)^{k_1}\cdots(\alpha_n)^{k_n}$. 因此

$$g(1, x_1, \cdots, x_n) = f(x_1, \cdots, x_n).$$

从而推出, 如果S由$\mathbb{P}^n$中方程式$g = 0$给定, 那么, $S \cap \mathbb{E}_0 = S_0$.

例1(有限截口) 以下我们默认$\mathfrak{K}=\mathbb{R}$.

1) 圆周S在射影齐次坐标中由方程式$\alpha_1^2+\alpha_2^2=\alpha_0^2$给定, 在图$\mathbb{E}_0$中它有方程式$x_1^2+x_2^2=1$. 它与相对于图$\mathbb{E}_0$的无穷远直线$\mathbb{P}(V_0)$的交集$S\cap\mathbb{P}(V_0)$由条件$\alpha_0=0$得出(这是$V_0$和$\mathbb{P}(V_0)$的方程式). 由$\alpha_1^2+\alpha_2^2=0$推出$\alpha_1=\alpha_2=0$. 这样的点($\alpha_0=\alpha_1=\alpha_2=0$)确实是不存在的, 所以, 这就意味着$S\cap\mathbb{P}(V_0)=\varnothing$.

2) 在图$\mathbb{E}_0$中以$x_1^2-x_2^2=1$为方程式的双曲线S的齐次坐标有方程式$\alpha_1^2-\alpha_2^2=\alpha_0$. 它与无穷远直线$\mathbb{P}(V_0)$的交集由条件$\alpha_0=0$, 也就是$\alpha_2=\pm\alpha_1$得到. 此时, $\alpha_1\neq 0$, 否则三个坐标就全都等于零了. 取定α_1, 交集的两个点可以写成$(0:1:1),(0:1:-1)$形式. 另一方面, 在图$\mathbb{E}_1$中, $\alpha_1\neq 0$, 从而在图中S的方程式可以设定为$x_1^2+x_2^2=1(x_1=\alpha_0/\alpha_1,\quad x_2=\alpha_2/\alpha_1)$, 即$S\cap\mathbb{E}_1$是圆周且$S\cap\mathbb{P}(V_1)=\varnothing$.

3) 抛物线$x_1=x_2^2$(在$\mathbb{E}_0$中的方程式)取$x_1=\alpha_1/\alpha_0,\ x_2=\alpha_2/\alpha_0$时可由方程式$\alpha_0\alpha_1=\alpha_2^2$给出. 与$\mathbb{P}(V_0)(\alpha_0=0)$的交只有一个点(二重根), 即$(0:1:0)$. 实现向另一个坐标系的转换, $\alpha_0=\beta_0-\beta_1,\alpha_1=\beta_0+\beta_1,\alpha_2=\beta_2$, 我们就得到$\beta_1^2+\beta_2^2=\beta_0^2$,它在新的图$\mathbb{E}_0'$中给出一个圆周. 这样一来, 圆周(或椭圆), 双曲线和抛物线只要在不同的仿射图上看, 它们在射影平面上都是同一种曲线. 特别地, 这个在解析几何学中熟知的结果被引到这里来仅仅就是给出所讨论的概念的一个实际例子.

6. 射影群 设$\mathbb{P}(V)$是域$\mathfrak{K}$上向量空间V生成的射影空间, 从而点$\tilde{\mathbf{x}}\in\mathbb{P}(V)$就是一条向量直线$\langle\mathbf{x}\rangle\subset V$. 设$\mathcal{A}:V\to V$是$V$上的一个非退化的线性算子. 它把直线变成直线, 且不会把直线变成$\mathbf{0}$. 可见, 下面的定义是有意义的.

定义4 *V上的每个线性算子$\mathcal{A}$都引导出一个变换$\tilde{\mathcal{A}}:\mathbb{P}(V)\to\mathbb{P}(V)$,*

$$\tilde{\mathcal{A}}\cdot\tilde{\mathbf{x}}=\widetilde{\mathcal{A}\mathbf{x}}, \tag{5}$$

*就称$\tilde{\mathcal{A}}$是个***射影变换**.

等式(5)与我们采用的点$\tilde{\mathbf{x}}$的定义是完全一致的, 所以,

$$\tilde{\mathcal{A}}\cdot\widetilde{\lambda\mathbf{x}}=\widetilde{\lambda\cdot\mathcal{A}\mathbf{x}}=\widetilde{\mathcal{A}\mathbf{x}}=\tilde{\mathcal{A}}\cdot\tilde{\mathbf{x}}. \tag{6}$$

由(6)可推出, 同样地, $\widetilde{\lambda\mathcal{A}}=\tilde{\mathcal{A}}$. 事实上, 可以相信

定理1 *等式$\tilde{\mathcal{B}}=\tilde{\mathcal{A}}$成立, 当且仅当, $\mathcal{B}=\lambda\mathcal{A}$.*

证明 我们只需指出, $\tilde{\mathcal{B}}=\tilde{\mathcal{A}}\Rightarrow\mathcal{B}=\lambda\mathcal{A}$. 因为, 对于$V$的任意一个向量$\mathbf{x}\neq\mathbf{0}$, $\widetilde{\mathcal{B}\mathbf{x}}=\tilde{\mathcal{B}}\tilde{\mathbf{x}}=\tilde{\mathcal{A}}\tilde{\mathbf{x}}=\widetilde{\mathcal{A}\mathbf{x}}$, 所以, 一定有某个由$\mathbf{x}$决定的一个纯量$\lambda_{\mathbf{x}}\neq 0$使得$\mathcal{B}\mathbf{x}=\lambda_{\mathbf{x}}\cdot\mathcal{A}\mathbf{x}$. 如果$\mathbf{y}=\alpha\mathbf{x}$, 那么

$$\lambda_{\mathbf{y}}\cdot\mathcal{A}\mathbf{y}=\mathcal{B}\mathbf{y}=\alpha\mathcal{B}\mathbf{x}=\alpha\lambda_{\mathbf{x}}\mathcal{A}\mathbf{x}=\lambda_{\mathbf{x}}\mathcal{A}\mathbf{y},$$

从而有$\lambda_{\mathbf{y}}=\lambda_{\mathbf{x}}$. 如果$\mathbf{x}$和$\mathbf{y}$是线性无关的向量, 那么, 向量$\mathcal{A}\mathbf{x}$和$\mathcal{A}\mathbf{y}$也将是线性无关的, 再由关系式

$$\begin{aligned}\lambda_{\mathbf{x}}\cdot\mathcal{A}\mathbf{x}+\lambda_{\mathbf{y}}\cdot\mathcal{A}\mathbf{y}&=\mathcal{B}\mathbf{x}+\mathcal{B}\mathbf{y}\\&=\mathcal{B}(\mathbf{x}+\mathbf{y})=\lambda_{\mathbf{x}+\mathbf{y}}\cdot\mathcal{A}(\mathbf{x}+\mathbf{y})=\lambda_{\mathbf{x}+\mathbf{y}}\mathcal{A}\mathbf{x}+\lambda_{\mathbf{x}+\mathbf{y}}\mathcal{A}\mathbf{y},\end{aligned}$$

就可以推出$\lambda_{\mathbf{x}} = \lambda_{\mathbf{x}+\mathbf{y}} = \lambda_{\mathbf{y}}$. 这意味着, $\lambda_{\mathbf{x}} = \lambda$是个不依赖于$\mathbf{x}$的纯量, 进而得到, $\mathcal{B} = \lambda\mathcal{A}$. □

为了获得射影变换$\tilde{\mathcal{A}}$ 的坐标记法, 我们在空间V 中选择一个基底$(\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_n)$并用矩阵$A = (a_{ij})$代表线性算子$\mathcal{A}$在这个基底之下的矩阵:

$$\mathcal{A}\mathbf{e}_j = \sum_{i=0}^{n} a_{ij}\mathbf{e}_i.$$

如果$\tilde{\mathbf{x}} = (\alpha_0 : \alpha_1 : \cdots : \alpha_n)$而$\tilde{\mathcal{A}}\tilde{\mathbf{x}} = (\beta_0 : \beta_1 : \cdots : \beta_n)$, 那么

$$\beta_i = \lambda \sum_{j=0}^{n} a_{ij}\alpha_j, \quad i = 0, 1, \cdots, n, \tag{7}$$

其中$\lambda \neq 0$是一个纯量. 这可以由等式

$$\lambda^{-1}\sum_{i=0}^{n} \beta_i \mathbf{e}_i = \mathcal{A}\mathbf{x} = \sum_{j=0}^{n} \alpha_j \mathcal{A}\mathbf{e}_j = \sum_{j=0}^{n}\sum_{i=0}^{n} a_{ij}\mathbf{e}_i$$

直接看出来, 而且与线性算子作用下坐标变换原则(见第2章)是一致的.

现在, 设$\mathbb{E}_0$是$\mathbb{P}(V)$中的仿射图. 它对于点$\tilde{\mathbf{x}} = (\alpha_0 : \alpha_1 : \cdots : \alpha_n)$, $\alpha_0 \neq 0$有意义. 如果在(7)中给出$\beta_0 \neq 0$, 那么, $\tilde{\mathcal{A}}\tilde{\mathbf{x}} \in \mathbb{E}_0$. $\tilde{\mathbf{x}}$的仿射坐标是$x_j = \alpha_j/\alpha_0$, $1 \leqslant j \leqslant n$, 而对于点$\tilde{\mathcal{A}}\tilde{\mathbf{x}}$的坐标是$y_j = \beta_j/\beta_0$, $1 \leqslant j \leqslant n$. 如果把标号为$i = 1, 2, \cdots, n$的等式(7)用$\beta_0$除之, 再注意到右侧的除数和分母$\alpha_0$, 我们就得到射影变换$\tilde{\mathcal{A}}$在图$\mathbb{E}_0$的仿射坐标之下的记法

$$y_i = \frac{a_{i1}x_1 + \cdots + a_{in}x_n + a_{i0}}{a_{01}x_1 + \cdots + a_{0n}x_n + a_{00}}, \quad 1 \leqslant i \leqslant n. \tag{8}$$

其特点, 别的不说, 在于所有这些公式的分母都是相同的. 不定因子λ也消失了.

说明 当然, 会有这样的形式, 使得射影变换$\tilde{\mathcal{A}}$完全可以把$\mathbb{E}_0$(精确地说, 集合$\Phi_0^{-1}(\mathbb{E}_0)$)的点变到不属于图的点, 也就是无穷远点(换句话说, 属于$\mathbb{P}(V_0)$). 在这种情形, 公式(8)就失去了意义. 如果$a_{00} = 1$; $a_{0j} = 0$, $1 \leqslant j \leqslant n$, 这个早就知道的事实就不会发生了. 可见, **仿射变换是射影变换的特殊情形**.

对于V中的任意向量$\mathbf{y} \neq \mathbf{0}$, 由于$\mathcal{A}$的非退化性, 必能找到这样一个向量$\mathbf{x}$使得$\mathcal{A}\mathbf{x} = \mathbf{y}$. 这意味着, 任意向量$\tilde{\mathbf{y}} \in \mathbb{P}(V)$必定是某个向量$\tilde{\mathbf{x}}$在射影变换$\tilde{\mathcal{A}}$之下的像: $\tilde{\mathbf{y}} = \widetilde{\mathcal{A}\mathbf{x}} = \tilde{\mathcal{A}}\tilde{\mathbf{x}}$. 类似地, 任意两个不同的点$\tilde{\mathbf{x}}$, $\tilde{\mathbf{z}}$必变成不同的点: $\tilde{\mathcal{A}}\tilde{\mathbf{x}} = \tilde{\mathcal{A}}\tilde{\mathbf{z}} \Rightarrow \widetilde{\mathcal{A}\mathbf{x}} = \widetilde{\mathcal{A}\mathbf{z}} \Rightarrow \mathcal{A}\mathbf{z} = \lambda\mathcal{A}\mathbf{x} \Rightarrow \mathcal{A}(\mathbf{z} - \lambda\mathbf{x}) = \mathbf{0} \Rightarrow \mathbf{z} = \lambda\mathbf{x} \Rightarrow \tilde{\mathbf{z}} = \tilde{\mathbf{x}}$. 换言之, 所有的射影变换都是双射.

定义5 在$\mathbb{P}(V)$到$\mathbb{P}(V)$的所有的双射变换构成的群中, 所有的射影变换构成一个子群, 用符号$PGL(V)$来代表, 并称之为**射影群**.

与记号$PGL(V)$紧紧相连的是, 这个群乃是完全线性群$GL(V)=GL_{n+1}(\mathfrak{K})$的一个同态像. 事实上, 映射$\Pi:\mathcal{A}\mapsto\tilde{\mathcal{A}}$满足同态条件$\pi(\mathcal{A}\mathcal{B})=\pi(\mathcal{A})\pi(\mathcal{B})$, 这是因为

$$\widetilde{\mathcal{A}\mathcal{B}}\tilde{\mathbf{x}}=\widetilde{\mathcal{A}\mathcal{B}\mathbf{x}}=\tilde{\mathcal{A}}\widetilde{\mathcal{B}\mathbf{x}}=\tilde{\mathcal{A}}\left(\tilde{\mathcal{B}}\tilde{\mathbf{x}}\right)=\left(\tilde{\mathcal{A}}\tilde{\mathcal{B}}\right)\tilde{\mathbf{x}}.$$

由定理1推出, 核Ker π由相似算子$\lambda\mathcal{E}$组成: $\tilde{\mathcal{A}}=\widetilde{\mathcal{E}}\Rightarrow\mathcal{A}=\lambda\mathcal{E}$. 因为, 映射$\Psi:\lambda\mapsto\lambda\mathcal{E}$显然是群$\mathfrak{K}^*$到Ker $\pi=\{\lambda\mathcal{E}\mid\lambda\in\mathfrak{K}^*\}$的同构, 所以, 在上述说法的基础上, 可得

定理2　空间$\mathbb{P}(V)$的所有射影变换构成射影群$PGL(V)$, 它是完全线性群$GL(V)$的一个同态像. 同态映射π的核同构于域$\mathfrak{K}$的乘法群$\mathfrak{K}^*$, 而且有短正合列

$$1\to\mathfrak{K}^*\xrightarrow{\Psi}GL(V)\xrightarrow{\pi}PGL(V)\to1.\tag{9}$$

在这个给定的情形, 就是不出现这个正合列, 事情也说得过去, 只要简单地知道Ψ是个同构嵌入, Im $\Psi=$ Ker π, 而π是个满射. 但是, 我们乘此机会为的是更加习惯于正合列概念的本身, 它被广泛地应用于现代数学.

7. 射影几何　我们知道, 射影群$PGL(V)$在$\mathbb{P}(V)$上作用有可迁性, 也就是它可以把任意一点变成任意一个另外的点. 与一般概念(见第4章§2的第4目)相对应, 群$PGL(V)$也对应一种几何学. 这个几何学就称为**射影几何学**. 射影几何学的课题是研究在$\mathbb{P}(V)$中的空间图形的, 在$PGL(V)$的变换作用下不变的那些性质. 这样的性质, 同样地, 被称为是**射影性质**. 直线的平行性和平面的平行性, 显然, 与射影性质没有关系. 毕达哥拉斯定理同样也不是一个射影性质, 因为它不包含长度和角度的概念, 我们需要的射影性质不仅将许多欧几里得几何学的定理排除在外, 也排除了仿射几何学的许多定理. 无论如何, 射影几何是一种内容极其丰富又是极其必要的几何学. 稍后, 我们将建立一个四点共线的重要射影性质. 而眼下, 我们先注意射影群$PGL(V)$的一系列性质.

i) 按照公式(8)后面的说明, 群$PGL(V)$包含作用在仿射图$\mathbb{E}_0$上的仿射群Aff($\mathbb{E}_0$)作为自己的一个子群(对于群Aff($\mathbb{E}_i$), $i=1,2,\cdots,n$也是一样).

ii) n维射影空间的点$\tilde{\mathbf{x}}_0,\tilde{\mathbf{x}}_1,\cdots,\tilde{\mathbf{x}}_{n+1}$可以说是位于一个**普通位置**, 如果它们中的任意$n+1$个点都不在一个超平面上. 换句话说, 任意$n+1$个向量

$$\mathbf{x}_0,\mathbf{x}_1,\cdots,\mathbf{x}_{i-1},\mathbf{x}_{i+1},\cdots,\mathbf{x}_{n+1}$$

线性无关.

定理3　设$\tilde{\mathbf{x}}_0,\tilde{\mathbf{x}}_1,\cdots,\tilde{\mathbf{x}}_{n+1}$和$\tilde{\mathbf{y}}_0,\tilde{\mathbf{y}}_1,\cdots,\tilde{\mathbf{y}}_{n+1}$是两个点组, 都是处于普通位置的. 那么, 必然存在, 而且是唯一的, 一个射影变换$\tilde{\mathcal{A}}\in PGL(V)$, 使得$\tilde{\mathcal{A}}\tilde{\mathbf{x}}_i=\tilde{\mathbf{y}}_i,\quad i=0,1,\cdots,n+1$.

证明　按照定义, 有

$$\langle\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_{n+1}\rangle=V=\langle\mathbf{y}_1,\mathbf{y}_2,\cdots,\mathbf{y}_{n+1}\rangle,$$

所以, 存在一个非退化的线性算子$\mathcal{A}'$, 使得

$$\mathcal{A}'\mathbf{x}_i = \mathbf{y}_i, \quad 1 \leqslant i \leqslant n+1. \tag{10}$$

乍一看, 这没有提供任何寻求$\mathcal{A}$的想法. 但是, 条件$\tilde{\mathcal{A}}\tilde{\mathbf{x}}_i = \tilde{\mathbf{y}}_i$用向量的术语来说, 对于$\mathbf{x}_i, \mathbf{y}_i \in V$就是要求能形如

$$\mathcal{A}\mathbf{x}_i = \lambda_i \mathbf{y}_i, \quad \lambda_i \neq 0, \quad 0 \leqslant i \leqslant n+1.$$

因为$\widetilde{\lambda_0^{-1}\mathcal{A}} = \tilde{\mathcal{A}}$, 即就是可以用条件$\lambda_0 = 1$来把$\mathcal{A}$规范化.

现在, 定义线性算子$\mathcal{B}$:

$$\mathcal{B}\mathbf{y}_i = \lambda_i \mathbf{y}_i, \quad 1 \leqslant i \leqslant n+1, \tag{11}$$

利用我们指定的这些λ_i, 从中挑选, 使得满足

$$\mathcal{B}\mathcal{A}'\mathbf{x}_0 = \mathbf{y}_0. \tag{12}$$

由这一组点处于普通位置的定义, 可得

$$\mathbf{x}_0 = \sum_{i=1}^{n+1} \alpha_i \mathbf{x}_i, \quad \mathbf{y}_0 = \sum_{i=1}^{n+1} \beta_i \mathbf{y}_i, \tag{13}$$

并且所有的系数α_i, β_i都不等于零. 与(10)相对应, 有

$$\mathcal{B}\mathcal{A}'\mathbf{x}_0 = \mathcal{B}\left(\sum_{i=1}^{n+1} \alpha_i \mathcal{A}'\mathbf{x}_i\right) = \mathcal{B}\left(\sum_{i=1}^{n+1} \alpha_i \mathbf{y}_i\right) = \sum_{i=1}^{n+1} \alpha_i \mathcal{B}\mathbf{y}_i = \sum_{i=1}^{n+1} \alpha_i \lambda_i \mathbf{y}_i.$$

借助于(13), 我们令$\alpha_i \lambda_i = \beta_i, \quad 1 \leqslant i \leqslant n+1$, 让它们满足(12). 纯量$\lambda_i = \alpha_i^{-1}\beta_i$以及和它们在一起的变换(11)就完全被确定了. 现在, 设

$$\mathcal{A} = \mathcal{B}\mathcal{A}',$$

我们就得到了一个唯一确定的线性算子, 它对应具有所有要求的性质的射影变换$\tilde{\mathcal{A}}$. □

显然, 定理3是第4章§3. 关于仿射变换的定理8的一个类似物.

推论 在射影直线$\mathbb{P}^1$上的两个三点组$\tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2$和$\tilde{\mathbf{y}}_0, \tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2$(每个组中的点均两两不同)上可唯一地定义一个$\mathbb{P}^1 \to \mathbb{P}^1$的射影变换, 它把$\tilde{\mathbf{x}}_i$分别变到$\tilde{\mathbf{y}}_i, i = 0, 1, 2$.

这个命题意味着, 对其余的情况, 一个点位于两个另外的点之间, 就不是一个射影性质. 由定理同样可以得到(关于$\tilde{\mathcal{A}}$的唯一性的命题), 并不是所有的在一条直线上的四个点都可以变成给定的点.

iii) 设$\mathbb{P}(U)$, $\mathbb{P}(W)$是$\mathbb{P}(V)$的两个m维平面. 那么, 它们必然是$PGL(V)$同余的, 也就是说, 它们能够在射影变换之下由一个变成另外一个.

事实上, 设

$$U = \langle \mathbf{u}_0, \mathbf{u}_1, \cdots, \mathbf{u}_m \rangle, \quad W = \langle \mathbf{w}_0, \mathbf{w}_1, \cdots, \mathbf{w}_m \rangle .$$

把$\{\mathbf{u}_0, \cdots, \mathbf{u}_m\}$扩充成空间$V$的基底$(\mathbf{u}_0, \cdots, \mathbf{u}_m, \cdots, \mathbf{u}_n)$, 把$\{\mathbf{w}_0, \cdots, \mathbf{w}_m\}$扩充成空间$V$的基底$(\mathbf{w}_0, \cdots, \mathbf{w}_m, \cdots, \mathbf{w}_n)$, 进而研究线性算子$\mathcal{A}$: $V \to V$, 使$\mathcal{A}\mathbf{u}_i = \mathbf{w}_i, i = 0, 1, \cdots, n$. 于是$\tilde{\mathcal{A}}\tilde{\mathbf{u}}_i = \tilde{\mathbf{w}}_i$, $\tilde{\mathcal{A}}\widetilde{\sum_i \alpha_i \mathbf{u}_i} = \mathcal{A}\widetilde{\sum_i \alpha_i \mathbf{u}_i} = \widetilde{\sum_i \alpha_i \mathbf{w}_i}$, 所以, $\tilde{\mathcal{A}}(\mathbb{P}(U)) = \mathbb{P}(W)$.

iv) *平面$\mathbb{P}(U) \subset \mathbb{P}(V)$上的所有的射影变换$\tilde{\mathcal{D}}$都可以开拓成整个空间$\mathbb{P}(V)$上的射影变换.*

实际上, 和$\mathbf{u}_0, \mathbf{u}_1, \cdots, \mathbf{u}_m$一起的向量$\mathcal{D}\mathbf{u}_0, \mathcal{D}\mathbf{u}_1, \cdots, \mathcal{D}\mathbf{u}_m$同样也构成子空间$U \subset V$的一个基底, 按前面说过的过程把它们扩充成$V$的基底

$$(\mathbf{u}_0, \cdots, \mathbf{u}_m, \mathbf{u}_{m+1}, \cdots, \mathbf{u}_n), \quad (\mathcal{D}\mathbf{u}_0, \cdots, \mathcal{D}\mathbf{u}_m, \mathbf{w}_{m+1}, \cdots, \mathbf{w}_n)$$

令

$$\mathcal{A}\mathbf{u}_i = \mathcal{D}\mathbf{u}_i, \quad 0 \leqslant i \leqslant m,$$
$$\mathcal{A}\mathbf{u}_i = \mathbf{w}_i, \quad m+1 \leqslant i \leqslant n,$$

我们就得到线性算子$\mathcal{A}: V \to V$, 它对应射影变换$\tilde{\mathcal{A}}$: $\mathbb{P}(V) \to \mathbb{P}(V)$, $\tilde{\mathcal{A}}$在$\mathbb{P}(U)$上与$\tilde{\mathcal{D}}$重合.

8. 重比(交比) 设$\mathbb{P}^n = \mathbb{P}(V)$是个射影空间, 而$\tilde{\mathbf{a}}_1$, $\tilde{\mathbf{a}}_2$, $\tilde{\mathbf{a}}_3$, $\tilde{\mathbf{a}}_4$ 是$\mathbb{P}(V)$ 中位于直线$\mathbb{P}^1 = \mathbb{P}(U)$ 上的四个点, 而且

$$\tilde{\mathbf{a}}_1 \neq \tilde{\mathbf{a}}_3, \quad \tilde{\mathbf{a}}_1 \neq \tilde{\mathbf{a}}_4, \quad \tilde{\mathbf{a}}_2 \neq \tilde{\mathbf{a}}_3, \quad \tilde{\mathbf{a}}_2 \neq \tilde{\mathbf{a}}_4.$$

这意味着

$$\langle \mathbf{a}_1, \mathbf{a}_3 \rangle = \langle \mathbf{a}_1, \mathbf{a}_4 \rangle = U = \langle \mathbf{a}_2, \mathbf{a}_3 \rangle = \langle \mathbf{a}_2, \mathbf{a}_4 \rangle .$$

我们把2维向量空间的基底$(\mathbf{c}, \mathbf{d})$向它的另外一个基底$(\mathbf{a}, \mathbf{b})$的转换矩阵

$$\mathbf{a} = \alpha\mathbf{c} + \beta\mathbf{d}, \quad \mathbf{b} = \gamma\mathbf{c} + \delta\mathbf{d}$$

的行列式记为

$$\left(\frac{\mathbf{a}, \mathbf{b}}{\mathbf{c}, \mathbf{d}}\right) = \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix}. \tag{14}$$

由已经加在点$\tilde{\mathbf{a}}_i$上的条件, 可以建立一个表达式

$$[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3, \tilde{\mathbf{a}}_4] = \left(\frac{\mathbf{a}_1, \mathbf{a}_3}{\mathbf{a}_1, \mathbf{a}_4}\right) \left(\frac{\mathbf{a}_2, \mathbf{a}_3}{\mathbf{a}_2, \mathbf{a}_4}\right)^{-1}. \tag{15}$$

定义6 *把表达式(15)称为四点$\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3, \tilde{\mathbf{a}}_4$的***重比(交比)**.

进而, 自然地, 可以验证, $[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3, \tilde{\mathbf{a}}_4]$仅仅与$\tilde{\mathbf{a}}_i$有关, 而与向量$\mathbf{a}_i$的选择没有关系(回忆一下, $\widetilde{\lambda \mathbf{a}_i} = \tilde{\mathbf{a}}_i$), 也就是说, 如果用$\lambda_i \mathbf{a}_i$来代替$\mathbf{a}_i$, $[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3, \tilde{\mathbf{a}}_4]$是不变的.

实际上, 我们用$\mathbf{b}_1 = \lambda \mathbf{a}_1$代替$\mathbf{a}_1$, 如果$\mathbf{a}_1 = \mathbf{a}_1, \mathbf{a}_3 = \gamma \mathbf{a}_1 + \delta \mathbf{a}_4$, 那么, $\mathbf{b}_1 = \mathbf{b}_1, \mathbf{a}_3 = \gamma \lambda^{-1} \mathbf{b}_1 + \delta \mathbf{a}_4$, 从而

$$\left(\frac{\mathbf{a}_1, \mathbf{a}_3}{\mathbf{a}_1, \mathbf{a}_4}\right) = \begin{vmatrix} 1 & 0 \\ \gamma & \delta \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ \gamma\lambda^{-1} & \delta \end{vmatrix} = \left(\frac{\mathbf{b}_1, \mathbf{a}_3}{\mathbf{b}_1, \mathbf{a}_4}\right).$$

此时, 显然, 因子

$$\left(\frac{\mathbf{a}_2, \mathbf{a}_3}{\mathbf{a}_2, \mathbf{a}_4}\right)^{-1}$$

也是不变的. 在用$\lambda \mathbf{a}_2$代替$\mathbf{a}_2$时, 可建立同样的事实.

现在, 用$\mathbf{b}_4 = \lambda \mathbf{a}_4$代替$\mathbf{a}_4$. 如果

$$\begin{aligned} &\mathbf{a}_1 = \mathbf{a}_1, && \mathbf{a}_2 = \mathbf{a}_2, \\ &\mathbf{a}_3 = \gamma \mathbf{a}_1 + \delta \mathbf{a}_4, && \mathbf{a}_3 = \gamma' \mathbf{a}_2 + \delta' \mathbf{a}_4, \end{aligned}$$

那么,

$$\begin{aligned} &\mathbf{a}_1 = \mathbf{a}_1, && \mathbf{a}_2 = \mathbf{a}_2, \\ &\mathbf{a}_3 = \gamma \mathbf{a}_1 + \delta\lambda^{-1} \mathbf{b}_4, && \mathbf{a}_3 = \gamma' \mathbf{a}_2 + \delta'\lambda^{-1} \mathbf{b}_4, \end{aligned}$$

进而就有

$$\left(\frac{\mathbf{a}_1, \mathbf{a}_3}{\mathbf{a}_1, \mathbf{a}_4}\right) = \begin{vmatrix} 1 & 0 \\ \gamma & \delta \end{vmatrix} = \lambda \begin{vmatrix} 1 & 0 \\ \gamma & \delta\lambda^{-1} \end{vmatrix} = \lambda \left(\frac{\mathbf{a}_1, \mathbf{b}_3}{\mathbf{a}_1, \mathbf{a}_4}\right),$$

$$\left(\frac{\mathbf{a}_2, \mathbf{a}_3}{\mathbf{a}_2, \mathbf{a}_4}\right) = \begin{vmatrix} 1 & 0 \\ \gamma' & \delta' \end{vmatrix} = \lambda \begin{vmatrix} 1 & 0 \\ \gamma' & \delta'\lambda^{-1} \end{vmatrix} = \lambda \left(\frac{\mathbf{a}_2, \mathbf{b}_3}{\mathbf{a}_2, \mathbf{a}_4}\right).$$

在此情况下, (15)右端的关系式没有变化. 当用$\lambda \mathbf{a}_3$替换$\mathbf{a}_3$时, 会得到同样的情况. 同时用$\lambda \mathbf{a}_i$代替$\mathbf{a}_i$可归纳为它们中的一连串的替换, 所以, 四个点的重比由公式(15)准确地确定.

定理4 *在射影变换之下重比不变, 即对任意$\tilde{\mathcal{A}} \in PGL(V)$都有*

$$\left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_3, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_4\right] = [\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3, \tilde{\mathbf{a}}_4]. \tag{16}$$

证明 设$\mathcal{A} \in PGL(V)$, 而U是V的一个2维子空间, $U' = \mathcal{A}(U)$. 由于$\mathcal{A}$的非退化性质, 任意两个线性无关的向量在$\mathcal{A}$的作用下仍然是线性无关的. 现在, 如果

$$\langle \mathbf{a}, \mathbf{b} \rangle = U = \langle \mathbf{c}, \mathbf{d} \rangle,$$

那么,

$$\langle \mathcal{A}\mathbf{a}, \mathcal{A}\mathbf{b} \rangle = U' = \langle \mathcal{A}\mathbf{c}, \mathcal{A}\mathbf{d} \rangle,$$

而联系到基底$(\mathbf{a},\mathbf{b})$和$(\mathbf{c},\mathbf{d})$的关系式:

$$\mathbf{a}=\alpha\mathbf{c}+\beta\mathbf{d},\quad \mathbf{b}=\gamma\mathbf{c}+\delta\mathbf{d},$$

在平面U'上对应同样的相等的关系式

$$\mathcal{A}\mathbf{a}=\alpha\mathcal{A}\mathbf{c}+\beta\mathcal{A}\mathbf{d},\quad \mathcal{A}\mathbf{b}=\gamma\mathcal{A}\mathbf{c}+\delta\mathcal{A}\mathbf{d}.$$

这意味着,

$$\left(\frac{\mathbf{a},\mathbf{b}}{\mathbf{c},\mathbf{d}}\right)=\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=\left(\frac{\mathcal{A}\mathbf{a},\mathcal{A}\mathbf{b}}{\mathcal{A}\mathbf{c},\mathcal{A}\mathbf{d}}\right).$$

于是, 用到我们的情形上, 就得到

$$\begin{aligned}\left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1,\tilde{\mathcal{A}}\tilde{\mathbf{a}}_2,\tilde{\mathcal{A}}\tilde{\mathbf{a}}_3,\tilde{\mathcal{A}}\tilde{\mathbf{a}}_4\right]&=\left[\widetilde{\mathcal{A}\mathbf{a}_1},\widetilde{\mathcal{A}\mathbf{a}_2},\widetilde{\mathcal{A}\mathbf{a}_3},\widetilde{\mathcal{A}\mathbf{a}_4}\right]\\ &=\left(\frac{\mathcal{A}\mathbf{a}_1,\mathcal{A}\mathbf{a}_3}{\mathcal{A}\mathbf{a}_1,\mathcal{A}\mathbf{a}_4}\right)\left(\frac{\mathcal{A}\mathbf{a}_2,\mathcal{A}\mathbf{a}_3}{\mathcal{A}\mathbf{a}_2,\mathcal{A}\mathbf{a}_4}\right)^{-1}=\left(\frac{\mathbf{a}_1,\mathbf{a}_3}{\mathbf{a}_1,\mathbf{a}_4}\right)\left(\frac{\mathbf{a}_2,\mathbf{a}_3}{\mathbf{a}_2,\mathbf{a}_4}\right)^{-1}\\ &=[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4].\end{aligned}$$

□

9. 重比的坐标表达式 设, 在直线$\mathbb{P}(U)$上有四个点$\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4$, 且对它们可以给出重比, 并且在$\mathbb{P}(U)$上引进一个坐标系. 也就是说, 设$U=\langle\mathbf{e},\mathbf{f}\rangle$且$\mathbf{a}_i=\alpha_i\mathbf{e}+\beta_i\mathbf{f}, i=1,2,3,4$. 那么

$$\tilde{\mathbf{a}}_i=(\alpha_i:\beta_i),\quad i=1,2,3,4. \tag{17}$$

(17)式的意义可以由射影空间的齐次坐标的一般定义中看得清清楚楚. 如果$\mathbf{a}_1=\mathbf{a}_1, \mathbf{a}_3=\gamma\mathbf{a}_1+\delta\mathbf{a}_4$, 那么

$$\left(\frac{\mathbf{a}_1,\mathbf{a}_3}{\mathbf{a}_1,\mathbf{a}_4}\right)=\delta. \tag{18}$$

此外, 由关系式

$$\alpha_3\mathbf{e}+\beta_3\mathbf{f}=\gamma(\alpha_1\mathbf{e}+\beta_1\mathbf{f})+\delta(\alpha_4\mathbf{e}+\beta_4\mathbf{f}),$$

我们得到$\alpha_3=\gamma\alpha_1+\delta\alpha_4, \beta_3=\gamma\beta_1+\delta\beta_4$. 所以,

$$\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_3 & \beta_3\end{vmatrix}=\begin{vmatrix}\alpha_1 & \beta_1\\ \gamma\alpha_1+\delta\alpha_4 & \gamma\beta_1+\delta\beta_4\end{vmatrix}=\begin{vmatrix}\alpha_1 & \beta_1\\ \delta\alpha_4 & \delta\beta_4\end{vmatrix}=\delta\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_4 & \beta_4\end{vmatrix},$$

再结合(18)式, 就给出了

$$\left(\frac{\mathbf{a}_1,\mathbf{a}_3}{\mathbf{a}_1,\mathbf{a}_4}\right)=\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_3 & \beta_3\end{vmatrix}\cdot\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_4 & \beta_4\end{vmatrix}^{-1}.$$

类似地, 有

$$\left(\frac{\mathbf{a}_2,\mathbf{a}_3}{\mathbf{a}_2,\mathbf{a}_4}\right)=\begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_3 & \beta_3\end{vmatrix}\cdot\begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_4 & \beta_4\end{vmatrix}^{-1}.$$

这样一来, 按照重比的定义6, 我们有

$$[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4]=\frac{\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_3 & \beta_3\end{vmatrix}\cdot\begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_4 & \beta_4\end{vmatrix}}{\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_4 & \beta_4\end{vmatrix}\cdot\begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_3 & \beta_3\end{vmatrix}}. \tag{19}$$

如果$\alpha_i\neq 0$, $1\leqslant i\leqslant 4$, 且$x_i=\beta_i/\alpha_i$, 那么, 由(19)式可以推出

$$[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4]=\frac{\begin{vmatrix}1 & x_1\\ 1 & x_3\end{vmatrix}\cdot\begin{vmatrix}1 & x_2\\ 1 & x_4\end{vmatrix}}{\begin{vmatrix}1 & x_1\\ 1 & x_4\end{vmatrix}\cdot\begin{vmatrix}1 & x_2\\ 1 & x_3\end{vmatrix}},$$

也就是

$$[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4]=\frac{(x_3-x_1)(x_4-x_2)}{(x_4-x_1)(x_3-x_2)}. \tag{20}$$

表达式(20)或者它的等价的表达式

$$[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4]=\frac{x_3-x_1}{x_2-x_3}:\frac{x_4-x_1}{x_2-x_4}$$

可以拿出来作为重比的定义, 但这不是很方便, 因为, 形式上它与仿射图的选择有关系(条件$\alpha_i\neq 0, 1\leqslant i\leqslant 4$), 这个时候, 作为表达式的等式(19)的右侧部分只有当点$\tilde{\mathbf{a}}_i$中有一个非正常的(无穷远的)才会有意义.

如果, 按照对于图$\mathbb{E}_0$的关系选择点$\tilde{\mathbf{a}}_4=(0:1)$作为无穷远点, 我们由(19)得到

$$[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4]=\frac{x_3-x_1}{x_3-x_2}.$$

除此之外, 如果还选择$\tilde{\mathbf{a}}_3$作为坐标原点$(1:0)$, 而选择$\tilde{\mathbf{a}}_2$作为单位点$(1:1)$, 那么, $x_3=0$, $x_2=1$, 且

$$[\tilde{\mathbf{a}}_1,\tilde{\mathbf{a}}_2,\tilde{\mathbf{a}}_3,\tilde{\mathbf{a}}_4]=x_1. \tag{21}$$

这就是在坐标系之下点$\tilde{\mathbf{a}}_1$的坐标, 在这个坐标系之下, $\tilde{\mathbf{a}}_4$是无穷远点, $\tilde{\mathbf{a}}_3$是零点, 而$\tilde{\mathbf{a}}_2$是单位点.

这个论断很有用, 我们不在$\mathbb{P}^1$=$\mathbb{P}(U)$中选取点, 而是选取一个坐标系. 设$\tilde{\mathbf{a}}_1$, $\tilde{\mathbf{a}}_2$, $\tilde{\mathbf{a}}_3$是直线$\mathbb{P}^1$上三个固定点, $\tilde{\mathbf{a}}_1\neq\tilde{\mathbf{a}}_2$. 令$\mathbf{e}=\nu\mathbf{a}_1$,$\mathbf{f}=\mu\mathbf{a}_2$,其中$\nu$, μ是纯量, 它们使得$\tilde{\mathbf{a}}_3=\mathbf{e}+\mathbf{f}$. 于是, $\tilde{\mathbf{a}}_1=(1:0)$, $\tilde{\mathbf{a}}_2=(0:1)$, $\tilde{\mathbf{a}}_3=(1:1)$. 现在, 如果$\tilde{\mathbf{a}}_4=(\alpha:\beta)$是$\mathbb{P}^1$上的任意一个点, 那么, 按着公式(19), 并且在其中取

$$\alpha_1=1,\beta_1=0;\quad \alpha_2=0,\beta_2=1;$$
$$\alpha_3=1,\beta_3=1;\quad \alpha_4=\alpha,\beta_4=\beta,$$

我们得到

$$[\widetilde{\mathbf{a}}_1, \widetilde{\mathbf{a}}_2, \widetilde{\mathbf{a}}_3, \widetilde{\mathbf{a}}_4] = \alpha/\beta.$$

可以看出, 点$\widetilde{\mathbf{a}}_4$给出的齐次坐标α, β的关系实际上是由重比$[\widetilde{\mathbf{a}}_1, \widetilde{\mathbf{a}}_2, \widetilde{\mathbf{a}}_3, \widetilde{\mathbf{a}}_4]$唯一确定的, 这样一来, 就有

定理5 *在射影直线$\mathbb{P}^1$上的两两不同的三个固定点$\widetilde{\mathbf{a}}_1$, $\widetilde{\mathbf{a}}_2$, $\widetilde{\mathbf{a}}_3$的情形下, 每个第四点$\widetilde{\mathbf{a}}_4 \in \mathbb{P}^1$都可以由重比$[\widetilde{\mathbf{a}}_1, \widetilde{\mathbf{a}}_2, \widetilde{\mathbf{a}}_3, \widetilde{\mathbf{a}}_4]$唯一确定.*

现在, 我们已经准备好了, 可以证明一个命题, 从本质上来讲, 细化了定理3的推论.

定理6 *在n维射影空间中, 两个分别共线的四点组$\widetilde{\mathbf{a}}_1$, $\widetilde{\mathbf{a}}_2$, $\widetilde{\mathbf{a}}_3$,$\widetilde{\mathbf{a}}_4$ 和$\tilde{\mathbf{b}}_1$, $\tilde{\mathbf{b}}_2$, $\tilde{\mathbf{b}}_3$, $\tilde{\mathbf{b}}_4$是$PGL(V)$同余的, 当且仅当,*

$$[\tilde{\mathbf{a}}_1, \widetilde{\mathbf{a}}_2, \widetilde{\mathbf{a}}_3, \widetilde{\mathbf{a}}_4] = \left[\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_3, \tilde{\mathbf{b}}_4\right].$$

证明 条件(22)的必要性, 一旦有$\tilde{\mathbf{b}}_i = \tilde{\mathcal{A}}\tilde{\mathbf{a}}_i$,由(16)即可推出.

现在, 设条件(22)已被满足. 按照射影群的性质3), 一定可以找到一个射影变换$\tilde{\mathcal{B}} : \mathbb{P}(V) \to \mathbb{P}(V)$, 它把各个点$\tilde{\mathbf{a}}_i$所在的直线$\mathbb{P}(U)$变到各个点$\tilde{\mathbf{b}}_i$所在的直线$\mathbb{P}(W)$. 再依据定理3的推论, 存在直线$\mathbb{P}(W)$的射影变换$\tilde{\mathcal{D}}$, 它把$\tilde{\mathcal{B}}\tilde{\mathbf{a}}_i$变成$\tilde{\mathbf{b}}_i$, $i = 1, 2, 3$. 利用射影群的性质4), 把$\tilde{\mathcal{D}}$开拓成整个空间$\mathbb{P}(V)$的射影变换$\tilde{\mathcal{A}}_1$. 变换$\tilde{\mathcal{A}} = \tilde{\mathcal{A}}_1\tilde{\mathcal{B}}$就把点$\tilde{\mathbf{a}}_i$变成了$\tilde{\mathbf{b}}_i$, $i = 1, 2, 3$. 对应地, 把点$\tilde{\mathbf{a}}_4$变成$\mathbb{P}(W)$上的某个点$\tilde{\mathbf{c}}_4$. 根据定理4, 我们有

$$[\tilde{\mathbf{a}}_1, \widetilde{\mathbf{a}}_2, \widetilde{\mathbf{a}}_3, \widetilde{\mathbf{a}}_4] = \left[\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_3, \tilde{\mathbf{c}}_4\right],$$

再考虑到(22), 就可以给出

$$\left[\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_3, \tilde{\mathbf{b}}_4\right] = \left[\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_3, \tilde{\mathbf{c}}_4\right].$$

由定理5可以推出, $\tilde{\mathbf{c}}_4$=$\tilde{\mathbf{b}}_4$. □

习 题

1. 在著名的长度为7的**汉明码**H中, 有射影平面$\mathbb{F}_2\mathbb{P}^2$上由7个点和7条直线构成的一个构形($7 = 2^2 + 2 + 1$). 这个构形可用图21表现之.

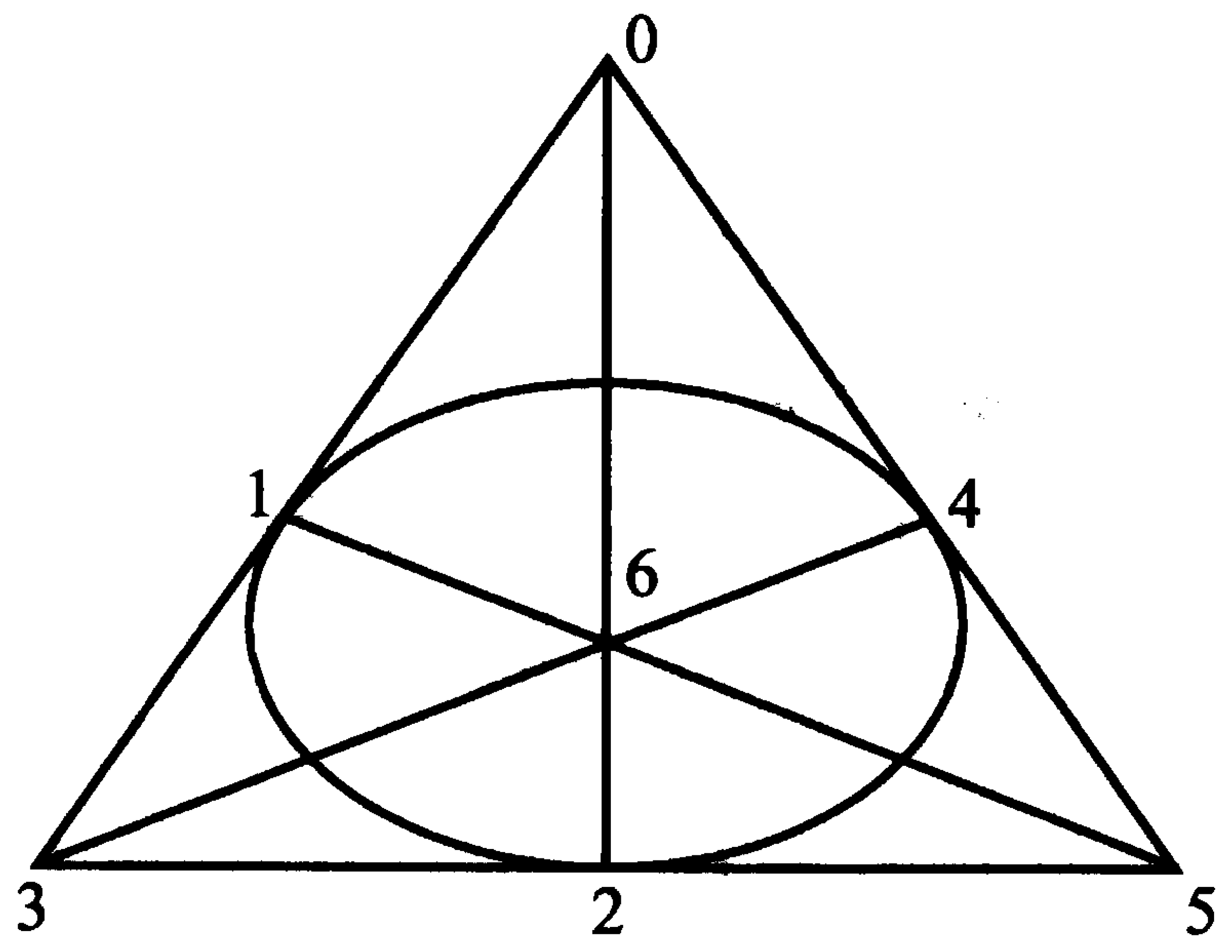


图21

关联矩阵J的每一行都有3个1: 它指明了在同一条直线上的3个点, 每一行同样也是一个H的权3的代码字. 而权4的代码字组成矩阵$\widetilde{J}$;

$$J=\begin{pmatrix}0&1&1&0&1&0&0\\0&0&1&1&0&1&0\\0&0&0&1&1&0&1\\1&0&0&0&1&1&0\\0&1&0&0&0&1&1\\1&0&1&0&0&0&1\\1&1&0&1&0&0&0\end{pmatrix},\widetilde{J}=\begin{pmatrix}1&0&0&1&0&1&1\\1&1&0&0&1&0&1\\1&1&1&0&0&1&0\\0&1&1&1&0&0&1\\1&0&1&1&1&0&0\\0&1&0&1&1&1&0\\0&0&1&0&1&1&1\end{pmatrix}.$$

矩阵$\widetilde{J}$是把矩阵J中的1换成0而把0换成1得到的. 在H中总共有16个码字, 我们已经写出了14个, 剩下的两个就是(0 0 0 0 0 0 0)和(1 1 1 1 1 1 1).

(1) 验证, 码H实际上就是$\mathbb{F}_2^7$的一个4维的子空间, 它是由线性方程组

$$x_1+x_3+x_4+x_5=0,\quad x_1+x_2+x_3+x_6=0,\quad x_2+x_3+x_4+x_7=0$$

决定的(在图22的3个圆中, 每个方程式启用位于同一个圆中的变量);

(2) 证明. $PGL(\mathbb{F}_2\mathbb{P}^2)$是个非交换的168阶的群, 它可以用作用在射影平面的点上的7阶置换群来实现;

(3) 证明

$$PGL(\mathbb{F}_2\mathbb{P}^2)=\mathrm{Aut}(H):=\{\,\sigma\in S_7\ |\sigma(H)=H\,\}.$$

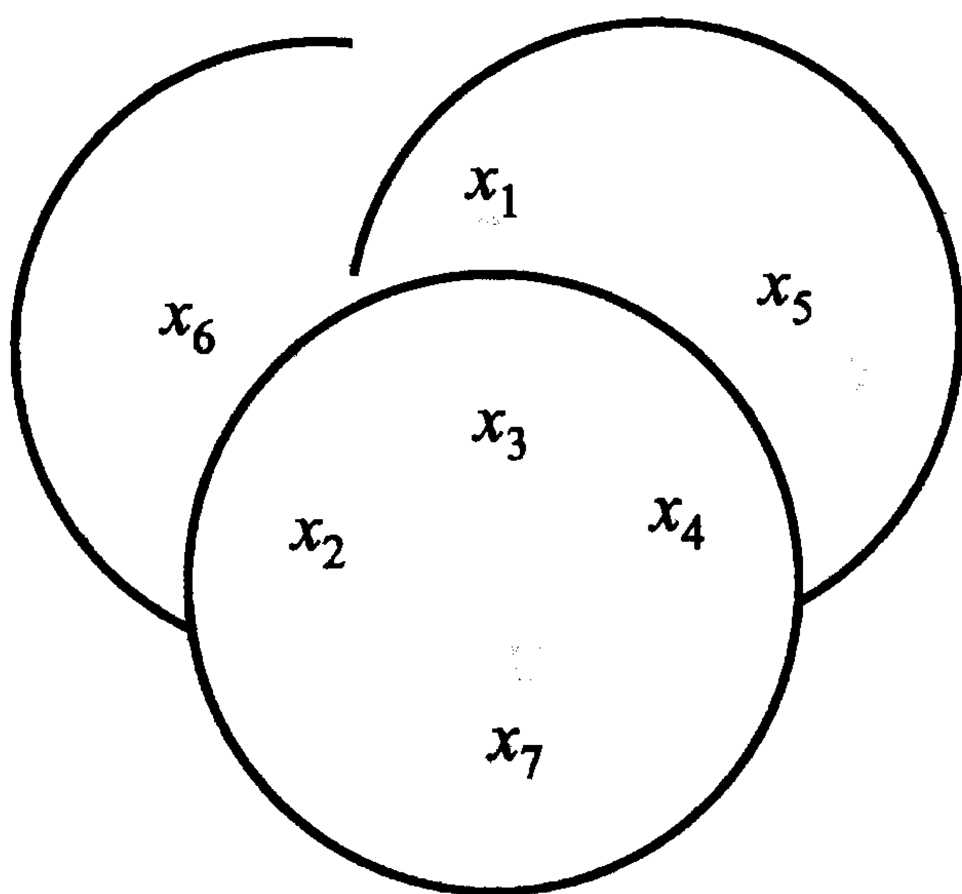


图22

2. 设$\Pi=\mathbb{F}_2\mathbb{P}^2$是个射影平面, 由7个点组成, 而$\tilde{p}$, $\tilde{q}$是平面Π上两个不同的点. 存在多少个自同构能把$\tilde{p}$变到$\tilde{q}$?

3. 证明, 在域$\mathbb{C}$上任意一个射影变换都至少有一个不动点.

§4 射影空间的二次曲面

1. 分类 设V是个$n+1\geqslant 2$的实向量空间. q是V上的二次型, f是与q配极的对称的双线性型, 所有满足方程式$q(\mathbf{x})$=0的向量$\mathbf{x}\in V$组成一个集合C, 可称它是以$\mathbf{0}$为顶点的(**逆向的**)锥. 如果这个集合不仅仅是由$\mathbf{0}$组成的, 那么, 在集合$V\setminus\{\mathbf{0}\}$到射影空间$\mathbb{P}(V)$的映射π之下$C\setminus\{\mathbf{0}\}$的像为$\widetilde{C}$(见§3), 即被称为**射影二次曲面**(当n=2时, 称

为射影二次曲线). 这个定义与前面给出的关于射影代数流形的一般定义是完全吻合的. $q(\mathbf{x})$只有在某个齐次坐标系之下才有意义. 如果二次型q是非退化的, 那么, 二次曲面$\tilde{C}$也说成是**非退化的**. 当$\tilde{C}$作为点的几何轨迹是$n-r$维, $r<n$的射影平面(具有条件$x_0^2+\cdots+x_r^2=0$)时, 如同在仿射情形一样, 就称$\tilde{C}$为**二重子空间**.

在V中任意选择一个基底$(\mathbf{e}_0,\mathbf{e}_1,\cdots,\mathbf{e}_n)$, 在相应的以$\mathbf{0}$为原点的坐标系之下锥$C$由形如

$$q(\mathbf{x})=\sum_{i,j=0}^{n}\varphi_{ij}x_ix_j=0 \tag{1}$$

的方程式给出. 方程式(1)也同样是二次曲面$\tilde{C}$在齐次坐标之下的方程式. 如所周知, 基底$(\mathbf{e}_i)$是可以用这样一种方式选取的, 即使得二次型q在这个基底之下取规范型. 可以借助于线性变换$\mathcal{A}\in GL(V)$完成从一个基底向另一个基底的转换, 而$\mathcal{A}$对应一个射影变换$\tilde{\mathcal{A}}\in PGL(V)$(见§3). 利用可将方程式两端乘以$-1$这样一条性质, 我们可以认为, 在$q$的规范型中正的平方项的个数不少于一半. 还需要回忆对于实二次型的惯性定律. 最终我们就得到下列命题.

定理1　*n维射影空间$\mathbb{P}(V)$中, 所有的非二重子空间的二次曲面$\tilde{C}$都等价于一个具有圆锥方程式*

$$x_0^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2=0,\quad [r/2]\leqslant s<r \tag{2}$$

的二次曲面.

为使$\mathbb{P}(V)$中的两个二次曲面射影等价, 必要而且只要, 它们的方程式的左侧具有相同的秩和相同的符号差(已假定指标满足$s\leqslant[r/2]$).

2. 射影二次曲面的例子与表现　我们看到, 射影二次曲面的分类与仿射几何以及欧几里得几何相比, 最为简单. §3的例1可以算得上是一个极好的解释: 仿射的不同的椭圆、抛物线和双曲线却在不同的仿射图中对应同一种射影曲线. 类似的情形, 在高维情况下也会出现. 如, 在$\mathbb{P}^3$中, 有两个不同的非退化的二次曲面, 它们分别由方程式

$$\alpha_0^2+\alpha_1^2+\alpha_2^2-\alpha_3^2=0,\quad \alpha_0^2+\alpha_1^2-\alpha_2^2-\alpha_3^2=0 \tag{3}$$

对应.

二次曲面$\tilde{C}$: $\alpha_0^2+\alpha_1^2+\alpha_2^2-\alpha_3^2=0$可引出椭球面、双叶双曲面和椭圆抛物面: 第一种情形对应图$\mathbb{E}_3=\mathbf{e}_3+V_3$, 第二种情形对应图$\mathbb{E}_0,\mathbb{E}_1$或$\mathbb{E}_2$, 而第三种情形的对应图在新的坐标系

$$\beta_0=\alpha_0,\quad \beta_1=\alpha_1,\quad \beta_2=\alpha_2-\alpha_3,\quad \beta_3=\alpha_2+\alpha_3$$

之下给出会更方便些, 此时, 二次曲面$\tilde{C}$的方程式可记成形如

$$\beta_0^2+\beta_1^2+\beta_2\beta_3=0,$$

它在图$\mathbb{E}' = \mathbf{e}_3' + V_3'$中对应椭圆抛物面. 在这三种情形下, 仿射二次曲面C_a与无穷远射影平面$\mathbb{P}(V_i)$的交集是有区别的:

$C_a \cap \mathbb{P}(V_3) = \varnothing$, 如果$C_a$是个椭球面;

$C_a \cap \mathbb{P}(V_2)$是个圆周, 如果C_a是双叶双曲面;

$C_a \cap \mathbb{P}(V_3')$是二重点$(0:0:1:0)$, 如果$C_a$是个椭球抛物面.

类似地可以说明, 对应(3)的第二个方程式的二次曲面可引出单叶双曲面和双曲抛物面, 不难得出一般情况下射影二次曲面在仿射图$(\mathbb{E}, \Phi)$上的公式(1)(非空的,而且不是二重子空间). 总能在空间V中选择一个基底$(\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_n)$使得双曲面$\mathbb{E}$和$\mathbb{E}_0 = \mathbf{e}_0 + V_0$重合. 于是$\Phi = \Phi_0$,且在坐标系$\{\mathbf{e}_0; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$之下,二次曲面$S = \tilde{C} \cap \mathbb{E} = \Phi_0(\tilde{C})$由方程式

$$Q(\mathbf{e}_0 + \mathbf{y}) = \sum_{i,j=1}^{n} \varphi_{ij} y_i y_j + 2\sum_{i=1}^{n} \varphi_{i0} y_i + \varphi_{00} \tag{4}$$

决定.

条件$(\varphi_{ij}) \neq 0$足够使以(4)为方程式的二次曲面S不为零,也不为二重子空间. 实际上,在相反的情况下,二次函数Q必然在某个坐标系$\{\mathbf{f}_0; \mathbf{f}_1, \cdots, \mathbf{f}_n\}$之下形如

$$\pm\left(\sum_{i=1}^{r} z_i^2 + \varepsilon\right),$$

其中$\varepsilon = 0$或$\varepsilon = 1$. 于是, 射影二次曲面在基底$(\mathbf{f}_i)$之下的齐次坐标必然是$\sum_{i=1}^{r} z_i^2 = 0$或者$\sum_{i=0}^{r} z_i^2 = 0$,但是,这不可能.

反之, 所有的仿射二次曲面S(非零的,且不为二重子空间, 所说的(4))都是$\mathbb{P}(V)$中一个射影二次曲面的精确的表达. 实际上,二次曲面S的方程式(4)在坐标系$\{\mathbf{e}_0; \mathbf{e}_1, \cdots, \mathbf{e}_n\}$之下对应形如(1)的二次函数$q$和圆锥$C$. 因为$(\varphi_{ij})_1^n \neq 0$,所以$(\varphi_{ij})_0^n \neq 0$. 其次, 如果在某个坐标之下$q(\mathbf{z}) = \pm\sum_{i=0}^{r} z_0^2$, 那么, C就是个平面,这是不可能的, 因为交集$S = C \cap \mathbb{E}$是个二次曲面. 换句话说,$q(\mathbf{x}) = 0$是射影二次曲面$\tilde{C}$的方程式且S是$\tilde{C}$在仿射图$\mathbb{E}$中的表现. S同时还是射影二次曲面$\widetilde{C}' \subset \mathbb{P}(V)$的表现, $\widetilde{C}'$的方程式是$q_1(\mathbf{x}) = 0$, 那么, 依据§2的定理1, 对应的二次函数Q和Q_1就应该是成比例的. 此时, q和q_1成比例,所以, $\widetilde{C} = \widetilde{C}'$.

把注意力集中到这样的一个事实上,即在射影空间里, 柱面和锥面之间的所有差别都被磨光了. 事实在于, 所有的柱面的直母线从仿射的观点来看都是平行的, 相交于无穷远点. 我们说, 3维的仿射空间中的锥面$x_1^2 + x_2^2 - x_3^2 = 0$, 椭圆柱面、抛物柱面和双曲柱面都不过是3维的射影空间中同一个柱面$\alpha^2 + \beta^2 - \gamma^2 = 0$的不同的仿射图的表达而已.

3. 直线与射影二次曲面的交 正如我们所期待的那样, 任意射影二次曲面与任意直线必有非空的交集(与仿射情形不同).

也就是说,设二次曲面$\tilde{C}$由方程式(1)给定, 而直线是形如

$$x_i = \mu a_i + \nu b_i, \quad i = 0, 1, \cdots, n \tag{5}$$

的方程式,其中$\tilde{\mathbf{a}} = (a_0 : a_1 : \cdots : a_n)$,$\tilde{\mathbf{b}} = (b_0 : b_1 : \cdots : b_n)$是在直线上的两个点. 把(5)代入(1)中之后,我们得到方程式

$$q(\mathbf{a})\mu^2 + 2f(\mathbf{a}, \mathbf{b})\mu\nu + q(\mathbf{b})\nu^2 = 0. \tag{6}$$

我们感兴趣的是解$(\mu, \nu) \neq (0, 0)$, 因为坐标$x_0, \cdots, x_n$不应当同时转化为零.

如果

$$D = f(\mathbf{a}, \mathbf{b})^2 - q(\mathbf{a})q(\mathbf{b}) \neq 0,$$

那么, 我们得到两个解, 对应直线与二次曲面$\tilde{C}$的两个不同的交点(可能是复共轭的).

如果$D = 0$, 但并不是所有的$q(\mathbf{a}), q(\mathbf{b}), f(\mathbf{a}, \mathbf{b})$都等于零, 那么, 直线与二次曲面有一个二重的共同点. 如果这一点$\tilde{p}$不是奇点, 即$\left.\dfrac{\partial q}{\partial x_i}\right|_{\tilde{p}} \neq 0$至少对某一个$i$成立, 那么, 在这一点直线与二次曲面相切.

最后, 当

$$q(\mathbf{a}) = q(\mathbf{b}) = f(\mathbf{a}, \mathbf{b}) = 0$$

时, 方程式(6)转化成一个恒等式, 且直线(5)整个地位于二次曲面$\tilde{C}$之中, 成为它的一条直母线.

我们看得出来, 渐近方向的概念对于射影二次曲面而言没有任何意义.

4. 关于射影二次曲面的一般说明 1)考察点$\tilde{\mathbf{a}} \in \mathbb{P}(V)$和方程式

$$f(\mathbf{a}, \mathbf{x}) = \sum_{i,j=0}^{n} \varphi_{ij} a_i x_j = 0, \tag{7}$$

其中, f是个双线性型, 与q配极. 如果所有的x_j的系数都能转化为零, 那么

$$\begin{gathered} \varphi_{00}a_0 + \varphi_{10}a_1 + \cdots + \varphi_{n0}a_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi_{0n}a_0 + \varphi_{1n}a_1 + \cdots + \varphi_{nn}a_n = 0, \end{gathered} \tag{8}$$

而且因为所有的齐次坐标a_i不同时为零, 故(8)等价于条件$\det(\varphi_{ij})_0^n = 0$. 这样一来, 只有在二次曲面$\tilde{C}$退化的情形($q(\mathbf{x}) = 0$), 表达式(7)才能是其上带有$\tilde{a}$的一个恒等式, 这时, $\tilde{a}$为一个奇点.

除去上面所研究过的情况, 我们看到方程(7)在$\mathbb{P}(V)$中确定了一个超曲面Π, 它称为点$\tilde{a}$关于二次曲面$\tilde{C}$的**配极**. 称点$\tilde{a}$是自己配极的Π的**极点**. 明显的, 只有当$\tilde{a} \in \tilde{C}$时, 点$\tilde{a}$的配极Π通过$\tilde{a}$.

任意一个双曲面Π:

$$p_0x_0 + p_1x_1 + \cdots + p_nx_n = 0$$

都有唯一的一个相对于任意一个非退化的二次曲面$\tilde{C} \subset \mathbb{P}(V)$ 的极点. 实际上, 在$\det(\varphi_{ij})_0^n \neq 0$ 的情形, 方程组

$$\varphi_{0j}a_0 + \varphi_{1j}a_1 + \cdots + \varphi_{nj}a_n = p_j, \quad 0 \leqslant j \leqslant n,$$

有而且只有一个解.

2) 二次曲面完全可以作为一个一般的代数流形自然地用复数的观点加以研究. 也就是说, 在$\mathbb{CP}^n$中的复的射影流形有可能给出(在没有实数解的情况下)更加精致化的拆分. 这就需要回顾有关实系数多项式根的问题. 在[BA I]中, 我们知道, 只有运用复数域$\mathbb{C}$, 才允许给出这一类问题的穷尽解. 在参考书[2]的第3部分引进了$\mathbb{RP}^n$和$\mathbb{CP}^n$的直观比较, 这个时候, 精髓在第3章§4的实数域$\mathbb{R}$上的射影空间$\mathbb{P}(V)$的复扩张$\mathbb{P}(V^{\mathbb{C}})$起重要作用. 标准嵌入$V \subset V^{\mathbb{C}}$允许把$V$中的每一条$\mathbb{R}$直线与它的复扩张——在$V^{\mathbb{C}}$中的$\mathbb{C}$直线看成是同一个东西. 这样, 就定义了嵌入$\mathbb{P}(V) \subset \mathbb{P}(V^{\mathbb{C}})$. 在$\mathbb{P}(V^{\mathbb{C}})$中的点就是实的射影空间$\mathbb{P}(V)$的“复点”.

习 题

1. 试验证, 在空间$\mathbb{RP}^3$中有八类射影等价的二次曲面(包括零二次曲面, 退化二次曲面以及二重子空间), 在适当的齐次坐标系之下, 这些等价类可由方程式

(1) $\xi_0^2 + \xi_1^2 + \xi_2^2 + \xi_3^2 = 0$;　(2) $\xi_0^2 + \xi_1^2 + \xi_2^2 - \xi_3^2 = 0$;

(3) $\xi_0^2 + \xi_1^2 - \xi_2^2 - \xi_3^2 = 0$;　(4) $\xi_0^2 + \xi_1^2 + \xi_2^2 = 0$;

(5) $\xi_0^2 + \xi_1^2 - \xi_2^2 = 0$;　(6) $\xi_0^2 + \xi_1^2 = 0$;

(7) $\xi_0^2 - \xi_1^2 = 0$;　(8) $\xi_0^2 = 0$

分别给出来.

2. 利用习题1和习题5.2.1, 指出(1)—(3)型的二次曲面的仿射表达.

第6章　张　　量

在整个的教学过程中, 我们需要和这样一些数学概念打交道, 如, 域$\mathfrak{K}$ 上的向量空间V, 与V 共轭(或对偶) 的线性函数空间V^*, V上的双线性型的空间$\mathcal{L}_2(V,\mathfrak{K})$, V上的线性算子空间$\mathcal{L}(V) := \mathcal{L}(V,V)$(经常用End V或者Hom(V, V)代表), 等等. 甚至, 还引进了半线性变换空间(见第1章§4的第1目). 照例, 以V, V^*, $\mathcal{L}(V,\mathfrak{K})$以及其他空间的元素的各种计算是在选择了基底且描述了(向量, 线性型和双线性型的)坐标向新的基底转换时的变换规律之后才能更有效地进行. 现在, 我们准备用统一的观点研究所有上述的概念并且导出某种形式体系或顺序上的计算, 这些东西产生于而且被成功地运用于物理学和几何学中.

§1　张量计算初步

1. 张量的概念　为了可以达到更广泛合理的概括性, 我们仅限于讨论某些特殊类型的多线性变换.

定义1　设$\mathfrak{K}$是个域, V是$\mathfrak{K}$上的向量空间, V^*是V的共轭空间, p和q都是大于等于零的整数,

$$V^p \times (V^*)^q = \underbrace{V \times \cdots \times V}_{p} \times \underbrace{V^* \times V^*}_{q}$$

是p重空间V和q重空间V^*的笛卡儿积. 所有的$p+q$重线性映射

$$f : V^p \times (V^*)^q \to \mathfrak{K}$$

都被称为是V上的(p,q)型, $p+q$价(或秩)的**张量**. 同样, 也说f是一个p次共变且q次反变的**混合张量**. 当$p=0$时, 就简单地说f是**反变的**, 而当$q=0$, 则说是**共变的** .

特殊的, (1, 0)型的张量就是通常的V上的线性函数, 也就是V^*的一个元素, 而(0, 1)型张量就是V^*上的一个线性函数, 也就是V^{**}的一个元素. 但是, 由于有限维空间的自反性(第1章§3的定理2), 在V和V^{**}之间存在自然同构, 允许把$\varphi \in V^{**}$与某个向量$\mathbf{x}_\varphi \in V$等同起来(或者把向量$\mathbf{x}$与线性函数$\varepsilon_\mathbf{x} \in V^{**}$等同起来). 这个等同可以在线性函数的记法

$$f(\mathbf{x}) = (f, \mathbf{x}) \tag{1}$$

之下实现, 这个记法前面已经使用过了. 当f固定时, 这是V上的一个线性函数, 而当$\mathbf{x}$固定时, 它就是V^*上的一个线性函数. 换言之, (0, 1)型张量可以认为是一个向量, 也就是V的一个元素.

进一步, (2, 0)型的共变张量显然是V上的一个双线性函数, 而(0, 2)型的反变张量就是V^*上的一个双线性函数.

最简单的混合型张量, 即(1, 1)型张量的解释是饶有兴趣的. 按定义$f(\mathbf{x}, u)$是个对$\mathbf{x} \in V$和对$u \in V^*$均有线性性的函数. 对任意固定的$\mathbf{x}$, 函数对u而言是线性的, 所以, 能够找到$\mathcal{F}\mathbf{x} \in V$, 使得

$$f(\mathbf{x},\ u) = (u,\ \mathcal{F}\mathbf{x}) \tag{2}$$

(我们利用了记法(1)). 因为$f(\alpha\mathbf{x} + \beta\mathbf{y}, u) = \alpha f(\mathbf{x}, u) + \beta f(\mathbf{y}, u)$, 所以

$$(u, \mathcal{F}(\alpha\mathbf{x} + \beta\mathbf{y})) = \alpha(u, \mathcal{F}\mathbf{x}) + \beta(u, \mathcal{F}\mathbf{y}) = (u, \alpha\mathcal{F}\mathbf{x} + \beta\mathcal{F}\mathbf{y}),$$

从而有

$$\mathcal{F}\,(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathcal{F}\mathbf{x} + \beta\mathcal{F}\mathbf{y},$$

也就是说, $\mathcal{F}$是V上的线性算子. 反过来, 对每个$\mathcal{F} \in \mathcal{L}(V)$, 按照公式(2)建立函数$f: V \times V^* \to \mathfrak{K}$, 它对于$\mathbf{x} \in V$和$u \in V^*$都是线性的. 明显地, $f \mapsto \mathcal{F}$是个双射. 这样一来, 每个(1, 1)型张量都对应而且是唯一确定的V上的一个线性算子.

还要约定, 把(0, 0)型张量理解为一个纯量(域$\mathfrak{K}$的元素). 我们可以得出这样的结论, 所有的秩$\leqslant 2$的张量, 对我们来说, 都是熟悉的.

V上所有(p, q)型张量的集合$\mathbb{T}^{p;q} = \mathbb{T}_p^q(V)$构成一个向量空间. 事实上, 如果$f, g \in \mathbb{T}_p^q(V)$且$\alpha, \beta \in \mathfrak{K}$, 那么, 自然地可以把$\alpha f + \beta g$理解为一个张量, 它由公式

$$\begin{aligned}&(\alpha f + \beta g)(\mathbf{v}_1, \cdots, \mathbf{v}_p; u_1, \cdots, u_q)\\ =\ &\alpha f(\mathbf{v}_1, \cdots, \mathbf{v}_p; u_1, \cdots, u_q) + \beta f(\mathbf{v}_1, \cdots, \mathbf{v}_p; u_1, \cdots, u_q)\end{aligned} \tag{3}$$

来定义.

2. 张量的乘积 首先, 设

$$f: V_1 \times \cdots \times V_r \to \mathfrak{K}, g: W_1 \times \cdots \times W_s \to \mathfrak{K}$$

是任意的多线性型. 这意味着, V_i, W_j是相互之间没有任何关系的向量空间.

定义2 把f和g的张量积理解为映射

$$f \otimes g : V_1 \times \cdots \times V_r \times W_1 \times \cdots \times W_s \to \mathfrak{K},$$

它用公式

$$(f \otimes g)(\mathbf{v}_1, \cdots, \mathbf{v}_r; \mathbf{w}_1, \cdots, \mathbf{w}_s) = f(\mathbf{v}_1, \cdots, \mathbf{v}_r) g(\mathbf{w}_1, \cdots, \mathbf{w}_s)$$

定义. 要从本质上强调, 变量$\mathbf{v}_i$和变量$\mathbf{w}_j$是没有关系的.

显然, $f \otimes g$是一个多线性映射, 所以, 可以打一个比喻, 在$\mathbf{v}_1, \cdots, \mathbf{v}_r; \mathbf{w}_2, \cdots, \mathbf{w}_s$固定时, 我们就得一个函数, 它与$g(\mathbf{w}_1, \cdots)$成比例, 对$\mathbf{w}_1$是线性的.

如果, 比如说, f和g是V上的线性函数, 那么

$$(f \otimes g)(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) g(\mathbf{y})$$

就是一个特殊类型的双线性函数, 而且, 这个平平常常的例子说明, 我们没有任何资本期待等式$f \otimes g = g \otimes f$. 事实上,

$$(g \otimes f)(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}) g(\mathbf{x}) \neq f(\mathbf{x}) g(\mathbf{y}) = (f \otimes g)(\mathbf{x}, \mathbf{y}).$$

同时, 对任意三个多线性函数(例如, $h : U_1 \times \cdots \times U_t \to \mathfrak{K}$), 结合律是满足的

$$(f \otimes g) \otimes h = f \otimes (g \otimes h). \tag{4}$$

这是因为, 作为函数, 左侧和右侧都与

$$f(\mathbf{v}_1, \cdots, \mathbf{v}_r) g(\mathbf{w}_1, \cdots, \mathbf{w}_s) h(\mathbf{u}_1, \cdots, \mathbf{u}_t)$$

重合.

现在设f是(p, q)型张量, g是(r, s)型张量, 那么$f \otimes g$就是在笛卡儿积

$$V^p \times (V^*)^q \times V^r \times (V^*)^s$$

上的多线性函数. 把这个笛卡儿积与

$$V^{p+r} \times (V^*)^{q+s}$$

等同起来. 对所有的$\mathbf{v}_i \in V$, $u_j \in V^*$, 用公式

$$\begin{aligned}&(f \otimes g)(\mathbf{v}_1, \cdots, \mathbf{v}_{p+r}; u_1, \cdots, u_{q+s}) \\ =& f(\mathbf{v}_1, \cdots, \mathbf{v}_p; u_1, \cdots, u_q) g(\mathbf{v}_{p+1}, \cdots, \mathbf{v}_{p+r}; u_{q+1}, \cdots, u_{q+s})\end{aligned} \tag{5}$$

加以定义, 我们可以把$f \otimes g$作为V上的$(p + r, q + s)$型的张量进行研究.

在下面, 用来区别不同类型变量的分号, 依照惯例, 将会省略掉.

定义3 称由公式(5)给定的张量$f \otimes g$是张量f和张量g的(**张量**)**乘积**.

例1 设f, g, h是V上的三个线性函数, 而$\mathbf{a}$和$\mathbf{b}$是V中的两个向量. 正如我们从前说明过的, 按已知的等同式, 可以把它们说成是三个(1, 0)型张量f, g, h和两个(0, 1)型张量$\mathbf{a}$, $\mathbf{b}$, 在这种情况下, 就可以把张量

$$t = f \otimes g \otimes h \otimes \mathbf{a} \otimes \mathbf{b}$$

说成是(3, 2)型的. 如果$\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in V, u, v \in V^*$, 那么,

$$t(\mathbf{x}, \mathbf{y}, \mathbf{z}, u, v) = f(\mathbf{x})g(\mathbf{y})h(\mathbf{z})\mathbf{a}(u)\mathbf{b}(v),$$

其中$\mathbf{a}(u)$, $\mathbf{b}(v)$应该理解为用公式(1)确定的表达式:$\mathbf{a}(u)=(\mathbf{a}, u)=(u, \mathbf{a})=u(\mathbf{a})$.

由公式(5)和张量线性组合$\alpha f + \beta g$的定义(3)可以看出, 张量的乘积具有分配律

$$\begin{aligned} &(\alpha f + \beta g) \otimes h = \alpha f \otimes h + \beta g \otimes h, \\ &h \otimes (\alpha f + \beta g) = \alpha h \otimes f + \beta h \otimes g. \end{aligned} \tag{6}$$

综上所述:

1) 对任意类型的张量定义了乘积运算$\otimes$;

2) 乘积的价等于各因子的价之和;

3) 张量乘积是结合的而且是分配的, 但不是交换的.

3. 张量的坐标 我们已经开始感觉到要清楚地区分V的元素和V^*的元素的必要性. 按照经典的理解, 张量分析开始于在空间$\mathbb{T}_p^q(V)$中选择基底, 并用自己的坐标去刻画张量. 通常, 在V和V^*选择(相互)对偶的基底

$$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle, \qquad V^* = \langle e^1, \cdots, e^n \rangle,$$

并且在基底的向量之间按上面说的办法排列指标. 为有直观性, 在我们这里, 指标在上方的向量采用平常的字体印刷. 在对应的坐标中, 指标的排列是"对立"的, 即, 我们设$\mathbf{x} \in V$, $\mathbf{x} = \sum\limits_i \alpha^i \mathbf{e}_i$, $f \in V^*$, $f = \sum\limits_j \beta_j e^j$. 回顾

$$(\mathbf{e}_i, e^j) = (e^j, \mathbf{e}_i) = \delta_i^j = \begin{cases} 0, i \neq j, \\ 1, i = j, \end{cases} \tag{7}$$

$$f(\mathbf{x}) = (f, \mathbf{x}) = \sum_i \alpha^i \beta_i$$

上方的指标绝对不可以与能够判断出来的方幂次数混淆起来. 不过, 在我们这里将不会遇到这种情况.

在张量分析中, 通常按所谓的**哑指标**来引入对称性, 我们在前面已经遇到它一次, 在后面还会遇到它一次. 所以, 一些经常接触张量的人, 就默认逐次做和, 而把求

和号约定可以省略. 例如, 把$\mathbf{x}=\alpha^i\mathbf{e}_i$和$\mathbf{x}=\sum\limits_i\alpha^i\mathbf{e}_i$等同起来. 我们不赞成这个约定, 但是, 约定按不同的指标求和可以用最简单的一个和号来代替, 即

$$\sum_i\sum_j\cdots\sum_k=\sum_{i,j,\cdots,k},$$

而且不申明求和的界限, 因为根据上下文, 很明显, 指标取什么值(通常由1到 $n=\dim V$).

设给定了一个(p,q)型张量T, 它的值可表达成

$$T^{j_1,\cdots,j_q}_{i_1,\cdots,i_p}:=T(\mathbf{e}_{i_1},\cdots,\mathbf{e}_{i_p},e^{j_1},\cdots,e^{j_q}) \tag{8}$$

的形式.

定义4 称数$T^{j_1,\cdots,j_q}_{i_1,\cdots,i_p}$是张量$T$在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$之下的**坐标**(**系数**, 或者**分量**).

我们赋予这个定义惯用的思想, 在(p,q)型张量的空间$\mathbb{T}^{p;q}$本身中选择一适当的基底, 也就是考察被称之为可分解的(p,q)型张量

$$e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes\mathbf{e}_{j_1}\otimes\cdots\otimes\mathbf{e}_{j_q}, \tag{9}$$

而且, 如同从前做过的一样, 把$\mathbf{e}_{j_1},\cdots,\mathbf{e}_{j_q}$与$V^*$上的线性函数$\mathbf{e}_j(f)=f(\mathbf{e}_j)=(f,\mathbf{e}_j)$等同起来. 因为$(e^i,\mathbf{e}_{i'})=\delta^i_{i'}$, $(\mathbf{e}_k,e^{k'})=\delta^{k'}_k$, 所以

$$(e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes\mathbf{e}_{j_1}\otimes\cdots\otimes\mathbf{e}_{j_q})(\mathbf{e}_{i'_1},\cdots,\mathbf{e}_{i'_p},e^{j'_1},\cdots,e^{j'_q})=\delta^{i_1}_{i'_1}\cdots\delta^{i_p}_{i'_p}\delta^{j'_1}_{j_1}\cdots\delta^{j'_q}_{j_q} \tag{10}$$

构作张量

$$T_1=\sum_{i,j}T^{j_1\cdots j_q}_{i_1\cdots i_p}e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes\mathbf{e}_{j_1}\otimes\cdots\otimes\mathbf{e}_{j_q},$$

它是张量(9)以(8)为系数的线性组合. 利用公式(3)和(10), 我们得到

$$T_1(\mathbf{e}_{i_1},\cdots,\mathbf{e}_{i_p},e^{j_1},\cdots,e^{j_q})=T^{j_1\cdots j_q}_{i_1\cdots i_p},$$

也就是说它恰恰是张量T的坐标(见(8)). 但是, 张量T的坐标是完全确定, 因为, 根据它的多重线性性质, 对任意向量

$$\mathbf{x}_1=\sum_{i_1}\xi^{i_1}\mathbf{e}_{i_1},\cdots,\mathbf{x}_p=\sum_{i_p}\rho^{i_p}\mathbf{e}_{i_p}$$

以及线性型

$$u^1=\sum_{j_1}\sigma_{j_1}e^{j_1},\cdots,u^q=\sum_{j_q}\tau_{j_q}e^{j_q}$$

有

$$T(\mathbf{x}_1,\cdots,\mathbf{x}_p,u^1,\cdots,u^q)=\sum_{i,j}T^{j_1\cdots j_q}_{i_1\cdots i_p}\xi^{i_1}\cdots\rho^{i_p}\sigma_{j_1}\cdots\tau_{j_q}. \tag{11}$$

在这个关系中只要我们回顾一下双线性型$f: V^2 \to \mathfrak{K}$, 并将它记成$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j} f_{ij}\, x_i\, y_j$. 唯一不同之处在于指标的标记方式. 现在, 我们把它记成$\sum f_{ij} x^i y^j$.

如果张量T_1和张量T的坐标重合, 张量本身就应当重合, 即

$$T = \sum_{i,j} T^{j_1 \cdots j_q}_{i_1 \cdots i_p} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes \mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_q}. \tag{12}$$

特别地, 每个双线性型f应形如

$$f = \sum_{ij} f_{ij} e^i \otimes e^j.$$

剩下来还需要指明的是, 对应不同的配套指标$i_1, \cdots, i_p, j_1, \cdots, j_q$, 型(9)的可分解张量是线性无关的. 这可以由计算它们值的规则(10)直接得到. 事实上, 假若存在一个线性关系式

$$\sum \lambda^{j_1 \cdots j_q}_{i_1 \cdots i_p} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes \mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_q} = 0,$$

它以$\lambda^{j_1 \cdots j_q}_{i_1 \cdots i_p} \in \mathfrak{K}$为系数, 即可转向这个等式的左侧, 如同前面处理张量T_1一样, 可以直接导出结论, $\lambda^{j_1 \cdots j_q}_{i_1 \cdots i_p} = 0$.

空间$\mathbb{T}^{p;q}$的维数等于(9)中不同的基底向量的个数. 从而它与$p+q$个指标$i_1, \cdots, i_p, j_1, \cdots, j_q$的所有不同的选择的个数$n^{p+q}$相同, 因为每个指标都可以不依赖其他指标从1取到$n$. 张量坐标(8)很容易理解的直观性使得可以把它排成一个空间立体表的形式. 立方体的维数等于T的价数. 我们熟知的行, 列, 方阵都是这种$(p+q)$维表的特殊情形.

将上面所讲的总结为下述论断.

定理1 *V上的(p, q)型张量构成一个n^{p+q}维的向量空间$\mathbb{T}^q_p(V)$, 它以*

$$e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes \mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_q}$$

构成一个基底, 其中$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是空间$V$的一个基底, 而$(e^1, \cdots, e^n)$是空间$V^*$的对偶的基底.

存在而且唯一地存在一个张量, 它具有预先给定的坐标$T^{j_1 \cdots j_q}_{i_1 \cdots i_p}$.

4. 在不同坐标系中的张量 我们需要得到一个向新的基底转化时张量坐标的替换规则. 设$(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$是空间$V$的另外一个基底, $(e'^1, \cdots, e'^n)$是V^*中与之对偶的基底. 而$A = (a^i_j)$是$(\mathbf{e}_i)$向$(\mathbf{e}'_i)$的转换矩阵. 矩阵$A = (a^i_j)$的元素是这样表达的, 上方的指标是指行数, 而下方的指标是指列数. 在这些表示中

$$\mathbf{e}'_k = \sum_i a^i_k \mathbf{e}_i, \quad k = 1, \cdots, n. \tag{13}$$

用$B = (b^i_j)$代表由(e^i)向(e'^i)的转换矩阵的转置矩阵, 我们应当记

$$e'^k = \sum_i b^k_i e^i. \tag{14}$$

这样才能按着"不同层次"同时表示指标的原则严格对应. 再引进一个辅助矩阵$B^{-1} = C = (c_j^i)$:

$$e^k = \sum_i c_i^k e'^i.$$

利用基底的对偶性质, 可以得到

$$c_j^k = (\sum_i c_i^k e'^i, \mathbf{e}'_j) = (e^k, \mathbf{e}'_j) = (e^k, \sum_i a_j^i \mathbf{e}_i) = a_j^k.$$

由此可见, $C = A$, 而且

$$e^k = \sum_i a_i^k e'^i.$$

其次, $B = A^{-1}$, 同时我们可以看出由(e^i)向(e'^i)的转换矩阵就是矩阵${}^tB = {}^t(A^{-1}) = {}^tA^{-1}$, 它被称为$A$的**转置逆矩阵**.

现在, 来求出我们的张量T在基底

$$e'^{i_1} \otimes \cdots \otimes e'^{i_p} \otimes \mathbf{e}'_{j_1} \otimes \cdots \otimes \mathbf{e}'_{j_q}$$

之下的坐标$T'^{j_1\cdots j_q}_{i_1\cdots i_p}$. 按定义(见(8)),

$$\begin{aligned} T &= \sum_{i,j} T^{j_1\cdots j_q}_{i_1\cdots i_p} e^{i_1} \otimes \cdots \otimes \mathbf{e}_{j_q} = \sum_{i,j} T'^{j'_1\cdots j'_q}_{i'_1\cdots i'_p} e'^{i'_1} \otimes \cdots \otimes \mathbf{e}'_{j'_q} \\ &= \sum (\sum b^{i'_1\cdots i'_p}_{i_1\cdots i_p} T'^{j'_1\cdots j'_q}_{i'_1\cdots i'_p} a^{j_1\cdots j_q}_{j'_1\cdots j'_q}) e^{i_1} \otimes \cdots \otimes \mathbf{e}_{j_q}. \end{aligned}$$

换言之, 下面的定理成立.

定理2 当由空间V和V^*的对偶的基底$(\mathbf{e}_i)$, (e^i)向相同的空间的新的对偶的基底按公式(13)和(14)转换时, (p,q)型张量T的坐标按公式

$$T^{j_1\cdots j_q}_{i_1\cdots i_p} = \sum_{i',j'} b^{i'_1\cdots i'_p}_{i_1\cdots i_p} T'^{j'_1\cdots j'_q}_{i'_1\cdots i'_p} a^{j_1\cdots j_q}_{j'_1\cdots j'_q} \tag{15}$$

转换.

通常说, 公式(15)中矩阵$A = (a_j^i)$在张量坐标的上指标起作用, 而矩阵$B = (b_j^i) = A^{-1}$在下指标上起作用.

在刚开始时给出的张量的定义现在可以换一种说法. 所谓V上的(p,q)型张量T, 是与空间的每一个基底联系在一起的一组n^{p+q}个纯量$T^{j_1\cdots j_q}_{i_1\cdots i_p}$(在下方有$p$个共变指标, 而在上方有$q$个反变指标), 使得在不同的基底之下对应的数组彼此之间用关系(15)联系起来.

前面在多重线性型语言上的陈述的张量的作用很容易在坐标基准上加以记叙. 比如说, S和T是同型的张量, 那么, 它们的**线性组合**$\alpha S + \beta T$可被称为是以

$$\alpha S^{j_1\cdots j_q}_{i_1\cdots i_p} + \beta T^{j_1\cdots j_q}_{i_1\cdots i_p}$$

的张量. **任意型张量**$(Q^{j_1\cdots j_q}_{i_1\cdots i_p})$和任意型张量$(R^{l_1\cdots l_t}_{k_1\cdots k_s})$的乘积张量的坐标是

$$T^{j_1\cdots j_q l_1\cdots l_t}_{i_1\cdots i_p k_1\cdots k_s} = Q^{j_1\cdots j_q}_{i_1\cdots i_p} R^{l_1\cdots l_t}_{k_1\cdots k_s}. \tag{16}$$

这个定义, 实质上, 与基底的选择没有关系, 因为把变换(15)用到(16)的左侧就会在右侧的因子Q和R之间引出成比例的量$a^{j'}_j, b^j_{j'}$.

例2 考察V上的线性算子$\mathcal{F}$. 我们已经看到(见例1), 可把$\mathcal{F}$解释成(1, 1)型张量. 这与矩阵表达是一致的. 如果$V=\langle \mathbf{e}_1,\cdots,\mathbf{e}_n\rangle$且$\mathcal{F}\mathbf{e}_k=\sum\limits_i f^i_k\mathbf{e}_i$, 那么, 按公式向新的基底

$$\mathbf{e}'_k=\sum_i a^i_k\mathbf{e}_i,\quad \mathbf{e}_k=\sum_i b^i_k\mathbf{e}'_i,\quad \sum_i b^i_k a^l_i=\delta^l_k$$

转化时, 将会有

$$\sum_s f'^s_k\mathbf{e}'_s=\mathcal{F}\mathbf{e}'_k=\sum_i a^i_k\mathcal{F}\mathbf{e}_i=\sum_{i,j}a^i_k f^j_i\mathbf{e}_j=\sum_{i,j,s}a^i_k f^j_i b^s_j\mathbf{e}'_s.$$

进而有

$$f'^s_k=\sum_{i,j}a^i_k f^j_i b^s_j, \tag{17}$$

这刚好应当对张量$F=(f^i_j)$一次共变加一次反变. 公式(17)的另一种表达我们在表达线性算子在不同基底之下的形式时已经遇见过了.

张量在数学和物理学中大部分作为服从于变换原则(15)的自然客体的几何对象. 此外, 对于物理学, 最重要的不是张量本身, 而是张量场(曲率张量, 引力场张量, 等等). 简单说来, 可以把张量场理解为一个从空间V到所有给定类型的张量的集合的一个映射. 详情可见 [2] 第4章§8.

5. 空间的张量积 张量的张量积运算允许自然地加以推广, 这在微分几何学, 群的表示理论和数学物理中有各式各样的应用. 不追求最大的概括性, 对我们来说未必有追求一般性的必要, (而如果要那样做, 这种引入的预先设立的目标是什么?)我们只限于有足够的兴趣去构造向量空间的张量积. 这个构造在[BA Ⅲ]中可以找到应用.

定理3 *设V, W是域$\mathfrak{K}$上的向量空间. 那么, 存在$\mathfrak{K}$上一个向量空间T和一个双线性映射$\tau: V\times W\to T$, 满足下列条件:*

(T1) *如果$\mathbf{v}_1,\cdots,\mathbf{v}_k\in V$是线性无关的,且$\mathbf{w}_1,\cdots,\mathbf{w}_k\in W$, 那么, $\sum\limits_{i=1}^k\tau(\mathbf{v}_i,\mathbf{w}_i)=0\Rightarrow\mathbf{w}_1=\mathbf{0},\cdots,\mathbf{w}_k=\mathbf{0}$;*

(T2) *如果$\mathbf{w}_1,\cdots,\mathbf{w}_k\in W$是线性无关的, 且$\mathbf{v}_1,\cdots,\mathbf{v}_k\in V$, 那么, $\sum\limits_{i=1}^k\tau(\mathbf{v}_i,\mathbf{w}_i)=0\Rightarrow\mathbf{v}_1=\mathbf{0},\cdots,\mathbf{v}_k=\mathbf{0}$;*

(T3) *τ是一个满的映射, 也就是*

$$T=\langle\tau(\mathbf{v},\mathbf{w})\,|\mathbf{v}\in V,\mathbf{w}\in W\rangle_{\mathfrak{K}}.$$

此外, 对于(τ, T)在如下的意义下是具有泛性的, 如果任意一个向量空间T'和任意一个双线性映射$\tau': V \times W \to T'$作成的对(τ', T'), 那么, 一定可以找到唯一的一个线性映射$\sigma: T \to T'$使得对任意$\mathbf{v} \in V, \mathbf{w} \in W$都有$\tau'(\mathbf{v}, \mathbf{w}) = \sigma(\tau(\mathbf{v}, \mathbf{w}))$.

证明 下面仅仅给出一些必要的论断的草稿.

i) 如果$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle_{\mathfrak{K}}, W = \langle \mathbf{f}_1, \cdots, \mathbf{f}_m \rangle_{\mathfrak{K}}$, 那么, (T1)—(T3)综合起来等价于唯一性条件:向量$\tau(\mathbf{e}_i, \mathbf{f}_j)$, $1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m$ 构成空间T的一个基底.

ii) 对于域$\mathfrak{K}$ 上任意一个mn 维向量空间T, 映射τ 和$\mathbf{v} = \sum_i \alpha_i \mathbf{e}_i$, $\mathbf{w} = \sum_j \beta_j \mathbf{f}_j$ 可以由关系式

$$\tau(\mathbf{v}, \mathbf{w}) = \sum_{i,j} \alpha_i \beta_j \mathbf{t}_{ij}$$

来定义. 其中$(\mathbf{t}_{ij} | 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m)$是$T$的基底. 按照1), 对$(\tau, T)$满足条件(T1)—(T3), 而且, 所有的对都可以用这种方式得到.

iii) 对于所有的具有一个双线性映射$\tau' : V \times W \to T'$的对$(\tau', T')$, 我们可定义一个线性映射$\sigma : T \to T'$, 令$\sigma(\sum \gamma_{ij} \mathbf{t}_{ij}) = \sum \gamma_{ij} \tau'(\mathbf{e}_i, \mathbf{f}_j), \gamma_{ij} \in \mathfrak{K}$. 依照i)和ii), $\tau'(\mathbf{v}, \mathbf{w}) = \sum \alpha_i \beta_j \tau'(\mathbf{e}_i, \mathbf{f}_j) = \sigma(\sum \alpha_i \beta_j \mathbf{t}_{ij}) = \sigma(\tau(\mathbf{v}, \mathbf{w}))$. 反过来:如果, $\sigma(\tau(\mathbf{v}, \mathbf{w})) = \tau'(\mathbf{v}, \mathbf{w})$, 那么, $\sigma(\mathbf{t}_{ij}) = \sigma(\tau(\mathbf{e}_i, \mathbf{f}_j)) = \tau'(\mathbf{e}_i, \mathbf{f}_j)$.

iv) 设, 存在两个都具有泛性的对(τ, T), (τ', T'), 我们容易发现$\sigma : T \to T'$和$\sigma' : T' \to T$实际上是互逆的同构映射: $\sigma' \circ \sigma = e_T$, $\sigma \circ \sigma' = e_{T'}$. 这样一来, $T \cong T'$, 同时, 同构映射$\sigma : T \to T'$具有本定理的叙述中的性质. □

定义5 称给定的空间V, W唯一确定的(精确到同构)对(τ, T)是这两个**向量空间的张量乘积**.

不难说明, 在预先设定$\dim T = nm$时, 系统(T1)—(T3)中的条件(T1)和(T2)是可以省略掉的, 剩下这三个条件中的一个条件就足以定义张量乘积了.

记$T = V \otimes_{\mathfrak{K}} W$或简单地记为$V \otimes W$, 我们还应当理解, 向量空间$T$已经附加了从笛卡儿积$V \times W$到$T$的映射$(\mathbf{v}, \mathbf{w}) \to \mathbf{v} \otimes \mathbf{w}$, 它满足(T1)—(T3). 换言之, 张量积$V \otimes W$的元素乃是有序对$\mathbf{v} \otimes \mathbf{w}$, $\mathbf{v} \in V$, $\mathbf{w} \in W$以$\mathfrak{K}$的元素为系数的形式上的线性组合. 这个时候, 被设计好了的下列条件被满足

$$\begin{aligned}
&(\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} - \mathbf{v}_1 \otimes \mathbf{w} - \mathbf{v}_2 \otimes \mathbf{w} = 0, \\
&\mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) - \mathbf{v} \otimes \mathbf{w}_1 - \mathbf{v} \otimes \mathbf{w}_2 = 0, \\
&\lambda \mathbf{v} \otimes \mathbf{w} - \mathbf{v} \otimes \lambda \mathbf{w} = 0, \ \lambda \in \mathfrak{K}, \\
&\lambda(\mathbf{v} \otimes \mathbf{w}) = \lambda \mathbf{v} \otimes \mathbf{w} = \mathbf{v} \otimes \lambda \mathbf{w}.
\end{aligned} \tag{18}$$

由定理3可以直接看出来, 双射映射$\mathbf{v} \otimes \mathbf{w} \mapsto \mathbf{w} \otimes \mathbf{v}$, $(\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w} \mapsto \mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w})$, $\mathbf{v} \otimes \lambda \mapsto \lambda \otimes \mathbf{v} \mapsto \lambda \mathbf{v}$可以建立起向量空间之间的同构,

$$\begin{aligned}
&V \otimes W \cong W \otimes V, \\
&(U \otimes V) \otimes W \cong U \otimes (V \otimes W),
\end{aligned}$$

$$V \otimes \mathfrak{K} \cong \mathfrak{K} \otimes V \cong V$$

可把它们称为**标准同构**(这些同构符号不可用等号代替). 同样, 分配律也是满足的

$$(U \oplus V) \otimes W \cong (U \otimes W) \oplus (V \otimes W),$$
$$U \otimes (V \oplus W) \cong (U \otimes V) \oplus (U \otimes W).$$

在研究结构时, 下面的举措应该是把空间和它上面的线性算子综合在一起.

定义6 设$\mathcal{A} : V \to V, \mathcal{B} : W \to W$是两个线性算子. 称线性算子

$$\mathcal{A} \otimes \mathcal{B} : V \otimes W \to V \otimes W,$$

是算子$\mathcal{A}$, $\mathcal{B}$的**张量积**, 它按规则

$$(\mathcal{A} \otimes \mathcal{B})(\mathbf{v} \otimes \mathbf{w}) = \mathcal{A}\mathbf{v} \otimes \mathcal{B}\mathbf{w} \tag{19}$$

起作用. (而且, 按照线性性质, $(\mathcal{A} \otimes \mathcal{B})(\sum(\mathbf{v}_i \otimes \mathbf{w}_i)) = \sum \mathcal{A}\mathbf{v}_i \otimes \mathcal{B}\mathbf{w}_i$).

明显地, 这个定义与关系式(18)是一致的. 例如

$$\begin{aligned}&\mathcal{A}(\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathcal{B}\mathbf{w} - \mathcal{A}\mathbf{v}_1 \otimes \mathcal{B}\mathbf{w} - \mathcal{A}\mathbf{v}_2 \otimes \mathcal{B}\mathbf{w}\\ = &(\mathcal{A}\mathbf{v}_1 + \mathcal{A}\mathbf{v}_2) \otimes \mathcal{B}\mathbf{w} - \mathcal{A}\mathbf{v}_1 \otimes \mathcal{B}\mathbf{w} - \mathcal{A}\mathbf{v}_2 \otimes \mathcal{B}\mathbf{w} = 0.\end{aligned}$$

所以, $\mathcal{A} \otimes \mathcal{B}$在$V \otimes W$上的作用是确切给定了的. 我们同样地注意到, 由(19)可直接计算出关系式

$$(\mathcal{A} \otimes \mathcal{B})(\mathcal{C} \otimes \mathcal{D}) = \mathcal{A}\mathcal{C} \otimes \mathcal{B}\mathcal{D},$$
$$(\mathcal{A} + \mathcal{C}) \otimes \mathcal{B} = \mathcal{A} \otimes \mathcal{B} + \mathcal{C} \otimes \mathcal{B},$$
$$\mathcal{A} \otimes (\mathcal{B} + \mathcal{D}) = \mathcal{A} \otimes \mathcal{B} + \mathcal{A} \otimes \mathcal{D},$$
$$\mathcal{A} \otimes \lambda\mathcal{B} = \lambda\mathcal{A} \otimes \mathcal{B} = \lambda(\mathcal{A} \otimes \mathcal{B}).$$

我们把它们留给读者去验证.

和从前一样, 设$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle$, $W = \langle \mathbf{f}_1, \cdots, \mathbf{f}_m \rangle$. 在基底

$$(\mathbf{e}_1 \otimes \mathbf{f}_1, \cdots, \mathbf{e}_1 \otimes \mathbf{f}_m, \cdots, \mathbf{e}_n \otimes \mathbf{f}_1, \cdots, \mathbf{e}_n \otimes \mathbf{f}_m)$$

之下, 算子$\mathcal{A} \otimes \mathcal{B}$的矩阵$A \otimes B$是$nm \times nm$维的. 我们记下

$$\mathcal{A}\mathbf{e}_i = \sum_{i'} \alpha_{i'i}\mathbf{e}_{i'}, \quad \mathcal{B}\mathbf{f}_j = \sum_{j'} \beta_{j'j}\mathbf{f}_{j'},$$

就得到

$$(\mathcal{A} \otimes \mathcal{B})(\mathbf{e}_i \otimes \mathbf{f}_j) = \sum_{i',j'} \alpha_{i'i}\beta_{j'j}\mathbf{e}_{i'} \otimes \mathbf{f}_{j'}.$$

这就是说, 由$A=(\alpha_{i'i}), B=(\beta_{j'j})$, 可以得到

$$A\otimes B=(\alpha_{i'i}\beta_{j'j})=\begin{pmatrix}\alpha_{11}B & \alpha_{12}B & \cdots & \alpha_{1n}B\\ \alpha_{21}B & \alpha_{22}B & \cdots & \alpha_{2n}B\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{n1}B & \alpha_{n2}B & \cdots & \alpha_{nn}B\end{pmatrix}. \tag{20}$$

特别地, 对于迹, 我们有公式

$$\mathrm{tr}A\otimes B=\alpha_{11}\mathrm{tr}B+\cdots+\alpha_{nn}\mathrm{tr}B=\mathrm{tr}A\cdot\mathrm{tr}B. \tag{21}$$

可以顺便指出

$$\begin{aligned}\det A\otimes B&=\det((A\otimes E_m)(E_n\otimes B))\\ &=\det(A\otimes E_m)\cdot\det(E_n\otimes B)=(\det A)^m(\det B)^n,\end{aligned} \tag{22}$$

所以, 线性算子$\mathcal{A}$和$\mathcal{B}$的非退化性可以导出它们的张量乘积$\mathcal{A}\otimes\mathcal{B}$的非退化性.

公式(21)和(22)在群表示论中经常使用.

习 题

1. **克罗内克符号**. 验证, δ_i^j是$\mathbb{T}_1^1(V)$的一个元素, 它代表空间V到自己的恒等映射.

2. **度量张量**. 设$V=\langle\mathbf{e}_1,\cdots,\mathbf{e}_m\rangle_{\mathbb{R}}$是带有纯量乘积的欧几里得空间, 由正定的双线性型

$$(\mathbf{x}\mid\mathbf{y})=g(\mathbf{x},\mathbf{y})=\sum_{i,j}g_{ij}x^ix^j,\quad g_{ij}=g(\mathbf{e}_i,\mathbf{e}_j)$$

给出这个纯量积(遵循在几何学中采用的惯例, 这里我们用字母g和符号g_{ij}). 通常, 称$G_0=(g_{ij})$是空间V的((2, 0)型的)**度量张量**. 这样一来, 度量张量的分量就是纯量积对于V的一个适当的基底的格拉姆矩阵的元素.

现在, 我们回忆一下与欧几里得空间V标准同构的对偶空间V^*, 且把V的基底$(\mathbf{e}_i)$和空间V^*的对偶基底$(\mathbf{e}^i)$等同起来:$(\mathbf{e}_i\ |*)=\mathbf{e}^i$(见第3章§1的(13)). 如果$g^{ij}=g(\mathbf{e}^i,\mathbf{e}^j)$, 那么, $G^0=(g^{ij})$同样地被称为是空间V的**度量张量**(现在已经是(0, 2)型了). 这个时候, g_{ij}和g^{ij}是度量张量G的共变坐标和反变坐标.

试验证, $G^0G_0=E$, 也就是

$$\sum_j g^{ij}g_{jk}=\delta_{k\cdot}^{i}.$$

3. **纯量域的提升**. 设V是个以$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$为基底的实向量空间, $V^{\mathbb{C}}$是它的复化空间(见第3章§4的第3目). 因为复数域$\mathbb{C}$是实数域$\mathbb{R}$上的向量空间. $(1,i)$是它的一个基底, 我们就可以构造一个向量空间

$$\mathbb{C}\otimes V=\langle 1\otimes\mathbf{e}_1,\cdots,1\otimes\mathbf{e}_n,i\otimes\mathbf{e}_1,\cdots,i\otimes\mathbf{e}_n\rangle.$$

试验证, $\mathbb{R}$线性映射

$$\mathbb{R}\otimes V\to V^{\mathbb{C}}:1\otimes\mathbf{e}_k\mapsto\mathbf{e}_k,\quad i\otimes\mathbf{e}_k\mapsto i\mathbf{e}_k$$

定义了$\mathbb{C}\otimes V$与$V^{\mathbb{C}}$之间的同构映射.

更一般地, 设$\mathfrak{K}$是域$\mathfrak{L}$的子域, V是域$\mathfrak{K}$上的向量空间. 开始时把$\mathfrak{L}$看成$\mathfrak{K}$上的一个向量空间, 我们构作张量积$\mathfrak{L}\otimes_{\mathfrak{K}} V$. 此后, 在其上引进$\mathfrak{L}$上的向量空间结构, 该结构中, 用$a\in\mathfrak{L}$做纯量乘法是

$$a(b\otimes\mathbf{x})=(ab)\otimes\mathbf{x},\quad a,b\in\mathfrak{L},\mathbf{x}\in V.$$

证明这个定义的适定性.

4. 设, $\dim V>1$, $\dim W>1$. 试说明, 在张量积$V\otimes W$中存在这样的元素, 它不能简单的表达成$\mathbf{v}\otimes\mathbf{w}$的形式.

5. 设A,B都是可逆方阵, 试给出$A\otimes B$的逆矩阵.

6. 把矩阵

$$\begin{pmatrix} \beta_{11}A & \beta_{12}A & \cdots & \beta_{1m}A \\ \beta_{21}A & \beta_{22}A & \cdots & \beta_{2m}A \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{m1}A & \beta_{m2}A & \cdots & \beta_{mm}A \end{pmatrix}$$

与分块矩阵(20)加以比较.

7. 对于阶数分别为n和m的带有复系数的方阵A,B, 利用它们可以被化成三角形矩阵的性质, 证明公式(22).

§2 张量的卷积, 对称化与交错化

1. 张量的卷积 设线性算子$\mathcal{F}:V\to V$ 在基底$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 之下的矩阵是$F=(f^i_j)$, 可以用不变的方式定义算子$\mathcal{F}$的迹(见第2章§2):

$$\mathrm{tr}\mathcal{F}=\sum_i f^i_i.$$

迹的不变性, 即它不依赖于基底的选择, 很容易从§1的公式(17)看出来. 把(f^i_j)解释成(1, 1)型的张量, 因为在基底$(\mathbf{e}'_i)$之下, 我们有

$$\sum_k f'^k_k=\sum_{i,j,k} a^i_k f^j_i b^k_j=\sum_{i,j} f^j_i\sum_k b^k_j a^i_k=\sum_{i,j} f^j_i\delta^i_j=\sum_i f^i_i=\mathrm{tr}\mathcal{F}.$$

在张量分析中, 一个与迹相关联的算子会对应更一般的对合算子. 为了给出它的定义, 最简单地, 暂时把$(p,\ q)$混合型张量T与作为一个$p+q$重的多线性型在任意向量$\mathbf{x}_1,\cdots,\mathbf{x}_p\in V$以及$u_1,\cdots,u_q\in V^*$处的值等同起来. 把除$\mathbf{x}_r$和$u_s$外的所有变量都固定, 我们得到多线性型

$$f(\mathbf{x}_r,u_s):=T(\cdots,\mathbf{x}_r,\cdots,u_s,\cdots).$$

定义1 称和

$$\overline{T}=\sum_k f(\mathbf{e}_k,e^k) \tag{1}$$

是张量T按共变r–指标和反变s–指标的**卷积**.

在和(1)中, 每个被加项$f(\mathbf{e}_k, e^k)$都是$\mathbf{x}_1, \cdots, \hat{\mathbf{x}}_r, \cdots, \hat{u}_s, \cdots, u_q$的多重线性型(一如既往, $\hat{a}$表示删掉符号a), 依赖于V的基底的选择, 但是$\overline{T}$不依赖于基底的选择. 实际上, 如果$\mathbf{e}'_k = \sum_i a^i_k \mathbf{e}_i$, $A = (a^i_k)$, 那么, $e'^k = \sum_i b^k_i e^i$, 其中$B = (b^k_i) = A^{-1}$ (见§1的第4目), 所以

$$\begin{aligned}\sum_k f(\mathbf{e}'_k, e'^k) &= \sum_{i,j,k} a^i_k b^k_j f(\mathbf{e}_i, e^j)\\ &= \sum_{i,j} (\sum_k b^k_j a^i_k) f(\mathbf{e}_i, e^j) = \sum_{i,j} \delta^i_j f(\mathbf{e}_i, e^j) = \sum_i f(\mathbf{e}_i, e^i) = \overline{T},\end{aligned}$$

这就证明了$\overline{T}$的不变性. 正如我们已经知道的, 双线性型f可以表达成

$$f(\mathbf{x}_r, u_s) = (u_s, \mathcal{F}\mathbf{x}_r)$$

的形式, 其中$\mathcal{F}$是依赖于$\mathbf{x}_1, \cdots, \hat{\mathbf{x}}_r, \cdots, \hat{u}_s, \cdots, u_q$的线性算子. 在这种情况下, $\overline{T} = \mathrm{tr}\mathcal{F}$, 而这又再一次地确认了$\overline{T}$与基底的选择的无关性. 并且给出了卷积算子和取迹的明显的联系.

如果用符号tr^s_r表示指标对r, s的卷积算子. 那么, tr^s_r就是一个

$$\mathbb{T}^q_p(V) \to \mathbb{T}^{q-1}_{p-1}(V)$$

的线性映射. 对于每个可分解的混合型张量

$$R = f_1 \otimes \cdots \otimes f_p \otimes \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_q, \quad \mathbf{v}_i \in V, \quad f_j \in V^*,$$

有

$$\mathrm{tr}^s_r(R) = (f_r, \mathbf{v}_s) f_1 \otimes \cdots \otimes \hat{f}_r \otimes \cdots \otimes \hat{\mathbf{v}}_s \otimes \cdots \otimes \mathbf{v}_q.$$

因为$(e^r, \mathbf{e}_s) = \delta^r_s$, 所以, 算子$\mathrm{tr}^s_r$可以很方便地表示成坐标形式. 设

$$T = \sum_{i,j} T^{j_1 \cdots j_q}_{i_1 \cdots i_p} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes \mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_q}$$

(可以与§1的公式(12)相比较). 那么,

$$\begin{aligned}\overline{T} = \mathrm{tr}^s_r(T) &= \sum_{i,j} T^{j_1 \cdots j_q}_{i_1 \cdots i_p} \mathrm{tr}^s_r(e^{i_1} \otimes \cdots \otimes \mathbf{e}_{j_q})\\ &= \sum_{i,j} T^{j_1 \cdots j_q}_{i_1 \cdots i_p} \delta^{i_r}_{j_s} e^{i_1} \otimes \cdots \otimes \hat{e}^{i_r} \otimes \cdots \otimes \hat{\mathbf{e}}_{j_s} \otimes \cdots \otimes \mathbf{e}_{j_q}\\ &= \sum_{i,j} T^{j_1 \cdots \hat{j}_s \cdots j_q}_{i_1 \cdots \hat{i}_r \cdots i_p} e^{i_1} \otimes \cdots \otimes \hat{e}^{\,i_r} \otimes \cdots \otimes \hat{\mathbf{e}}_{j_s} \otimes \cdots \otimes \mathbf{e}_{j_q},\end{aligned}$$

其中

$$\overline{T}^{j_1 \cdots j_{s-1} j_{s+1} \cdots j_q}_{i_1 \cdots i_{r-1} i_{r+1} \cdots i_p} = \sum_k T^{j_1 \cdots j_{s-1} k j_{s+1} \cdots j_q}_{i_1 \cdots i_{r-1} k i_{r+1} \cdots i_p}. \tag{2}$$

我们知道, 可分解张量

$$e^{i_1}\otimes\cdots\otimes\hat{e}^{i_r}\otimes\cdots\otimes e^{i_p}\otimes\mathbf{e}_{j_1}\otimes\cdots\otimes\hat{\mathbf{e}}_{j_s}\otimes\cdots\otimes\mathbf{e}_{j_q}$$

组成空间$\mathbb{T}_{p-1}^{q-1}(V)$的一个基底, 因此, $T_{i_1\cdots i_{r-1}i_{r+1}\cdots i_p}^{j_1\cdots j_{s-1}j_{s+1}\cdots j_q}$就是张量$T$的坐标. 我们已经证明了如下的

定理1 *(p,q)混合型张量T按r–共变指标和s–反变指标的卷积是个$(p-1,q-1)$型的张量$\overline{T}$, 它的坐标由公式(2)确定.*

可以把卷积运算用到张量$\overline{T}=\mathrm{tr}_r^s(T)$本身上, 做$m$重卷积后, 就得到一个纯量或者得到纯共变(或者纯反变张量), 就不能再在它上面做卷积了, 这里m=min(p, q). 在这种情况. 就称它是个**完整卷积**. 取线性算子的迹就是一个完整卷积. 两个线性算子$\mathcal{A}$和$\mathcal{B}$的乘积可以作为一个不完整卷积的例子. 事实上, 如果$A=(a_j^i), B=(b_l^k)$是它们在某个基底之下分别对应的矩阵, 那么, 张量

$$T=T_{jl}^{ik},\quad T_{jl}^{ik}=a_j^ib_l^k,$$

按张量A的共变指标和张量B的反变指标的卷积就等于张量$C=(c_l^i)$, 其中

$$c_l^i=\sum_j T_{jl}^{ij}=\sum_j a_j^ib_l^j.$$

很容易猜想到, c_l^i是矩阵AB的元素.

研究过的例子刻画了一个经常应用的运算, 可用它构作两个张量(不同时都为共变的也不同时为反变的)的张量乘积以及构作随后由得到的混合型张量按一对或若干对指标做卷积.

2. 结构张量代数 设V是域$\mathfrak{K}$上的一个有限维代数(见第2章§2的定义1), 它的乘法运算$(a,b)\mapsto a*b$不一定是结合的. 如果$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$是空间$V$的一个基底, 那么,

$$\mathbf{e}_i*\mathbf{e}_j=\sum_k\gamma_{ij}^k\mathbf{e}_k$$

定义2 *称纯量$\gamma_{ij}^k\in\mathfrak{K}$是代数$V$在给定的基底之下的***结构常数**.

由于运算的双线性性质, 代数V的乘法运算$*$就由给定的基底以及这个基底之下的结构常数完全确定了. 但是, 在另外一个基底$(\mathbf{e}_1',\cdots,\mathbf{e}_n')$之下的结构常数就是另外的

$$\mathbf{e}_i'*\mathbf{e}_j'=\sum_k{\gamma'}_{ij}^k\mathbf{e}_k',$$

而且, 我们打算找出在不同的基底之下的结构常数之间的关系. 设

$$\mathbf{e}_i'=\sum_s a_i^s\mathbf{e}_s,\quad \mathbf{e}_j=\sum_t b_j^t\mathbf{e}_t',$$

则$B = (b_j^t) = A^{-1}, A = (a_i^s)$. 我们有

$$\sum_k {\gamma'}_{ij}^k bf{e'}_k = {\mathbf{e}'}_i * {\mathbf{e}'}_j = \left(\sum_s a_i^s \mathbf{e}_s\right) * \left(\sum_j^t \mathbf{e}_t\right)$$
$$= \sum_{s,t} a_i^s a_j^t \mathbf{e}_s * \mathbf{e}_t = \sum_{s,t,r} a_i^s a_j^t \gamma_{st}^r \mathbf{e}_r = \sum_{s,t,r,k} a_i^s a_j^t \gamma_{st}^r b_r^k \mathbf{e}'_k,$$

从而

$${\gamma'}_{ij}^k = \sum_{s,t,r} a_i^s a_j^t \gamma_{st}^r b_r^k.$$

把这个公式与§1的公式(15)加以比较, 我们看到, 结构常数可以替换成(2, 1)型张量的坐标. 因而就理由去考察代数V的所说的结构张量$\Gamma = (\gamma_{ij}^k)$. 代数V的Γ的值是完全确定的.

对一个固定的$\mathbf{a} \in V$, 研究映射$L_\mathbf{a} : \mathbf{x} \mapsto \mathbf{a} * \mathbf{x}$, 显然, 它是$V$上的一个线性算子. 对于研究代数$V$的结构, 迹形式——双线性对称型:

$$f_V(\mathbf{a}, \mathbf{b}) = \mathrm{tr} L_\mathbf{a} L_\mathbf{b} \tag{3}$$

是一个重要的工具. 型f_V的双线性性质可以从运算$*$的双线性性推出, 它的对称性则可由一般性质$\mathrm{tr}\mathcal{AB} = \mathrm{tr}\ \mathcal{BA}$推导出来(第2章§2的关系式(12′)).

为了显易地计算f_V, 我们记

$$\mathbf{a} = \sum_i \alpha^i \mathbf{e}_i, \quad \mathbf{b} = \sum_i \beta^i \mathbf{e}_i,$$
$$L_\mathbf{a} L_\mathbf{b} \mathbf{e}_k = \mathbf{a} * (\mathbf{b} * \mathbf{e}_k) = \sum_{i,j} \alpha^i \beta^j \mathbf{e}_i * (\mathbf{e}_j * \mathbf{e}_k) = \sum_{i,j,s} \alpha^i \beta^j \gamma_{jk}^s \gamma_{is}^t \mathbf{e}_t.$$

要计算$f_V(\mathbf{a}, \mathbf{b})$, 需要取矩阵

$$(R_k^t), R_k^t = \sum_{i,j,s} \alpha^i \beta^j \gamma_{jk}^s \gamma_{is}^t$$

的对角线上的元素且把它们加到一起:

$$f_V(\mathbf{a}, \mathbf{b}) = \sum_{i,j,s,t} \alpha^i \beta^j \gamma_{jk}^s \gamma_{is}^t. \tag{4}$$

正如我们所期待的一样, 我们已经把$f_V(\mathbf{a}, \mathbf{b})$记成张量的完整的卷积形式了.

例1 从力学和物理学的观点看, 研究3维李代数$V = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle$是有意义的(见第2章§2的例6), 它的乘积运算$(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} * \mathbf{b} = [\mathbf{a}, \mathbf{b}]$按3维的欧几里得空间中向量间的乘法给出

$$[\mathbf{e}_1, \mathbf{e}_2] = \mathbf{e}_3, \quad [\mathbf{e}_3, \mathbf{e}_1] = \mathbf{e}_2, \quad [\mathbf{e}_2, \mathbf{e}_3] = \mathbf{e}_1.$$

这里, 只有当三个指标i, j, k两两不同时才会有$\gamma_{ij}^k \neq 0$. 而且$\gamma_{ji}^k = -\gamma_{ij}^k$. 对任意向量

$$\mathbf{a} = \alpha^1\mathbf{e}_1 + \alpha^2\mathbf{e}_2 + \alpha^3\mathbf{e}_3, \quad \mathbf{b} = \beta^1\mathbf{e}_1 + \beta^2\mathbf{e}_2 + \beta^3\mathbf{e}_3,$$

按通常的公式来计算纯量乘积

$$(\mathbf{a}\,|\mathbf{b}) = \alpha^1\beta^1 + \alpha^2\beta^2 + \alpha^3\beta^3, \tag{5}$$

而按照公式

$$[\mathbf{a}, \mathbf{b}] = (\alpha^2\beta^3 - \alpha^3\beta^2)\mathbf{e}_1 + (\alpha^3\beta^1 - \alpha^1\beta^3)\mathbf{e}_2 + (\alpha^1\beta^2 - \alpha^2\beta^1)\mathbf{e}_3 \tag{6}$$

计算代数中的乘积. 我们应该记起在这种联系中坐标表达式是个关键.

运用公式(6), 可以直接验证, 纯量乘积(5)具有"结合性"

$$([\mathbf{a}, \mathbf{b}]\,|\mathbf{c}) = (\mathbf{a}|\,[\mathbf{b}, \mathbf{c}]). \tag{7}$$

这个漂亮的关系式实际上是在任意一个有限维李代数V中迹形式的可结合性

$$f_V([\mathbf{a}, \mathbf{b}]\,, \mathbf{c}) = f_V(\mathbf{a}, [\mathbf{b}, \mathbf{c}]) \tag{8}$$

的一个推广. 为了证明(8), 我们需要注意雅可比恒等式

$$[[\mathbf{x}, \mathbf{y}]\,, \mathbf{z}] + [[\mathbf{z}, \mathbf{x}]\,, \mathbf{y}] + [[\mathbf{y}, \mathbf{z}]\,\mathbf{x}] = 0.$$

它又可写成

$$[[\mathbf{x}, \mathbf{y}]\,, \mathbf{z}] = [\mathbf{x}, [\mathbf{y}, \mathbf{z}]] - [\mathbf{y}, [\mathbf{x}, \mathbf{z}]]\,,$$

于是, 对任意$\mathbf{x}, \mathbf{y} \in V$, 有

$$L_{[\mathbf{x},\mathbf{y}]} = L_\mathbf{x}L_\mathbf{y} - L_\mathbf{y}L_\mathbf{x}.$$

因为$\mathrm{tr}\ \mathcal{AB} = \mathrm{tr}\ \mathcal{BA}$, 所以, 只要设$\mathcal{A} = L_\mathbf{b}L_\mathbf{a}, \mathcal{B} = L_\mathbf{c}$或者$\mathcal{A} = L_\mathbf{c}L_\mathbf{b}, \mathcal{B} = L_\mathbf{a}$, 我们就得到

$$\begin{aligned} f_V([\mathbf{a}, \mathbf{b}], \mathbf{c}) &= \mathrm{tr}L_{[\mathbf{a},\mathbf{b}]}L_\mathbf{c} = \mathrm{tr}(L_\mathbf{a}L_\mathbf{b}L_\mathbf{c} - L_\mathbf{b}L_\mathbf{a}L_\mathbf{c}) \\ &= \mathrm{tr}L_\mathbf{a}L_\mathbf{b}L_\mathbf{c} - \mathrm{tr}L_\mathbf{c}L_\mathbf{b}L_\mathbf{a} = \mathrm{tr}L_\mathbf{a}L_\mathbf{b}L_\mathbf{c} - \mathrm{tr}L_\mathbf{a}L_\mathbf{c}L_\mathbf{b} \\ &= \mathrm{tr}L_\mathbf{a}(L_\mathbf{b}L_\mathbf{c} - L_\mathbf{c}L_\mathbf{b}) = \mathrm{tr}L_\mathbf{a}L_{[\mathbf{b},\mathbf{c}]} = f_V(\mathbf{a}, [\mathbf{b}, \mathbf{c}]), \end{aligned}$$

这就给出了关系式(8).

在3维代数的情形, (7)和(8)的等价性可以直接按张量公式(4)计算$f_V(\mathbf{a}, \mathbf{b})$ 的得数推出来

$$f_V(\mathbf{a}, \mathbf{b}) = -2(\alpha^1\beta^1 + \alpha^2\beta^2 + \alpha^3\beta^3) = -2(\mathbf{a}\,|\mathbf{b}).$$

再来注意下面的情况. 公式(7)改写成

$$(L_\mathbf{a}\mathbf{x}\,|\mathbf{y}) + (\mathbf{x}\,|L_\mathbf{a}\mathbf{y}) = 0$$

的形式, 就简单而又直接地表明了算子$L_{\mathbf{a}}$在任意一个基底之下所对应的矩阵的斜对称性, 它也可以被理解为算子$\mathcal{A} \in O_3(\mathbb{R}) = \mathrm{Aut}((\mathbf{x}|\mathbf{y}))$:

$$(\mathcal{A}\mathbf{x}\,|\,\mathcal{A}\mathbf{y}) = (\mathbf{x}\,|\,\mathbf{y}).$$

原来, 从这里可以最直接地延伸到一般群论和李代数理论.

3. 对称张量 在双线性型理论中, 我们把精力集中在两大类上:对称型和斜对称型. 在张量的情形, 可以按选出来的联立的共变或反变的指标集合来讨论对称性和斜对称性. 例如, 按最前边的两个共变指标和按最后边的两个反变指标对称就简单地意味着有等式

$$T^{r\cdots st}_{ij\cdots k} = T^{r\cdots ts}_{ij\cdots k} = T^{r\cdots st}_{ji\cdots k} = T^{r\cdots ts}_{ji\cdots k}.$$

用反变指标置换共变指标, 一般地说, 不能产生出有意义的工作量也不能导出向量.

讨论对称与斜对称问题, 如果我们限制在$(p,\,0)$型或者$(0,\,q)$型张量, 并且把置换应用到所有p个(或q个)指标上, 而不是只用到它们的一部分, 我们就一点也不会减少一般性. 此外, 域$\mathfrak{K}$, 到现在为止, 总是默认为有零特征数. 在应用中, 最重要的域是实数域$\mathbb{R}$和复数域$\mathbb{C}$.

换言之, 设对于一个给定的$T \in \mathbb{T}^0_p(V)$, 也就是

$$T = \sum_{i_1\cdots i_p} T_{i_1\cdots i_p} e^{i_1} \otimes \cdots \otimes e^{i_p}, \tag{9}$$

而S_p是作用在指标集合$\{1, 2, \cdots, p\}$上的p阶对称群. 对任意置换$\pi \in S_p$, 我们设

$$f_\pi(T)(\mathbf{x}_1, \cdots, \mathbf{x}_p) = T(\mathbf{x}_{\pi 1}, \cdots, \mathbf{x}_{\pi p}) \tag{10}$$

(这里, $\mathbf{x}_i$是以i为指标的向量;它的第k个坐标就是x_i^k).

因为T是V^p上的多线性型, 所以, $f_\pi(T)$就是个多线性型, 同时也是$(p,\,0)$型的. 事实上, 比方说, $\pi k = 1$, 那么

$$\begin{aligned} f_\pi(T)(\alpha\mathbf{x}_1 + \beta\mathbf{y}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p) &= T(\mathbf{x}_{\pi 1}, \cdots, \alpha\mathbf{x}_1 + \beta\mathbf{y}_1, \cdots, \mathbf{x}_{\pi p}) \\ &= \alpha T(\mathbf{x}_{\pi 1}, \cdots, \mathbf{x}_1, \cdots, \mathbf{x}_{\pi p}) + \beta T(\mathbf{x}_{\pi 1}, \cdots, \mathbf{y}_1, \cdots, \mathbf{x}_{\pi p}) \\ &= \alpha f_\pi(T)(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p) + \beta f_\pi(T)(\mathbf{y}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p). \end{aligned}$$

还可以按定义设

$$f_\pi(\alpha T' + \beta T'') = \alpha f_\pi(T') + \beta f_\pi(T''),$$

我们看得出来, π引导出一个非退化的线性算子$f_\pi : \mathbb{T}^{p;0} \to \mathbb{T}^{p;0}$. 进一步容易看出, $f_\sigma \circ f_\tau = f_{\sigma\tau}$(见[BA I]第1章§8的第4目).

与张量T经基底元素$e^{i_1} \otimes \cdots \otimes e^{i_p}$的表达式(9)相对应, 它的坐标是$T_{i_1\cdots i_p} = T(\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_p})$. 而张量$f_\pi T$在同样的基底之下的坐标是

$$f_\pi(T)_{i_1\cdots i_p} = f_\pi(T)(\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_p}) = T(\mathbf{e}_{i_{\pi 1}}, \cdots, \mathbf{e}_{i_{\pi p}}) = T_{i_{\pi 1}\cdots i_{\pi p}},$$

也就是

$$f_\pi(T)=\sum_{i_1\cdots i_p} T_{i_{\pi 1}\cdots i_{\pi p}} e^{i_1}\otimes\cdots\otimes e^{i_p},$$

或者, 等价于

$$f_\pi(T)=\sum_{i_1\cdots i_p} T_{i_1\cdots i_p} e^{i_{\pi^{-1}1}}\otimes\cdots\otimes e^{i_{\pi^{-1}p}}. \tag{11}$$

从而, 也顺便得到了

$$f_{\pi^{-1}}(e^{i_1}\otimes\cdots\otimes e^{i_p})=e^{i_{\pi 1}}\otimes\cdots\otimes e^{i_{\pi p}}.$$

当然, 如果我们要研究S_p作用在反变张量上的情形, 那么, 也就能够得到

$$f_{\pi^{-1}}(T)=\sum_{i_1\cdots i_p} T^{i_1\cdots i_p}\mathbf{e}_{i_{\pi^{-1}1}}\otimes\cdots\otimes\mathbf{e}_{i_{\pi^{-1}p}}. \tag{11'}$$

定义3 称$(p, 0)$型的张量T (或者$(0, q)$型)是**对称的**, 如果, 对每个$\pi\in S_p$都有 $f_\pi(T)=T$ (或者, 相应地, 对每个$\pi\in S_q$). 称映射

$$S=\frac{1}{p!}\sum_{\pi\in S_p} f_\pi: T_p^0(V)\to T_p^0(V) \tag{12}$$

为$T_p^0(V)$中的向量的**对称化**.

在$\mathbb{T}_p^0(V)$和$\mathbb{T}_0^q(V)$中的对称张量的子空间分别用$\mathbb{T}_p^+(V)$和$\mathbb{T}_+^q(V)$代表.

例如,

$$S(e^1\otimes e^2\otimes e^2)=\frac{1}{3}(e^1\otimes e^2\otimes e^2+e^2\otimes e^1\otimes e^2+e^2\otimes e^2\otimes e^1).$$

显然, $\mathbb{T}_p^0$的每个张量经对称化之后都是对称的

$$f_\sigma(S(T))=\frac{1}{p!}\sum_{\pi\in S_p} f_\sigma(f_\pi(T))=\frac{1}{p!}\sum_{\pi\in S_p} f_{\sigma\pi}(T)=\frac{1}{p!}\sum_{\tau\in S_p} f_\tau(T)=S(T).$$

我们这里要用到如下的事实, 对于固定的置换$\sigma\in S_p$, 随着π取遍群S_p的所有元素, $\sigma\pi$也取遍所有的元素, 而且只取一次. 换言之, Im $S\subset\mathbb{T}_p^+(V)$.

反过来, 在对称的张量上施对称化是恒等算子. 因为这可以由(12)直接推出来, 即$T\in\mathbb{T}_p^+(V)\Rightarrow T=S(T)$. 可见, 我们有

定理2 对称化映射S在$\mathbb{T}_p^0$上的作用具有性质$S^2=S$且Im$S=\mathbb{T}_p^+(V)$.

例2 张量的一个古典例子是**惯性张量**——3阶的对称矩阵$J=(J_{ij})$, 其中J_{ii}是刚体的能量相对于轴$\mathbf{e}_i$的轴矩, 而J_{ij}, $i\neq j$是带反号的离心转动惯量. 换句话说, 设给定了一个以$\dot{o}$为固定点的旋转刚体. 假设刚体由n个质量分别为m_k的质点组成;这

些质点的位置用向量行$[x_k, y_k, z_k]$, $1 \leqslant k \leqslant n$给出. 在计算刚体的角矩和动能时要用到的刻画质量分布的张量J可以由矩阵关系式

$$J = (\sum_k (x_k, y_k, z_k)[x_k, y_k, z_k])E - \sum_k [x_k, y_k, z_k](x_k, y_k, z_k) \qquad (*)$$

给出. (这里, 照惯例, E是3阶的单位矩阵, $(x_k, y_k, z_k) = {}^t[x_k, y_k, z_k]$是行向量;在刚体质量非连续分布的情形, 用求和代替求积分). 借助正交矩阵A, 按通常的规则$[x_k, y_k, z_k] \to A[x_k, y_k, z_k] = [x'_k, y'_k, z'_k]$(显然, A在质量m_k上不起作用), 我们得到矩阵

$$J' = AJ^tA = AJA^{-1},$$

也就是$J'_{ij} = \sum_{r,s} a_i^r a_j^s J_{rs}$, 这应当认为是个2价的张量. 如果把分量$J_{ij}$写成明显形式, 那么, 由(*)就有

$$J = \begin{pmatrix} \sum\limits_k m_k(y_k^2 + z_k^2) & -\sum\limits_k m_k x_k y_k & -\sum\limits_k m_k x_k z_k \\ -\sum\limits_k m_k x_k y_k & \sum\limits_k m_k(x_k^2 + z_k^2) & -\sum\limits_k m_k y_k z_k \\ -\sum\limits_k m_k x_k z_k & -\sum\limits_k m_k y_k z_k & \sum\limits_k m_k(x_k^2 + y_k^2) \end{pmatrix}$$

坐标J_{ij}不能被看成是具有独立意义的, 不依赖于坐标架选择的物理量, 但是, 正如我们看到的, 在整个J中获得了这样的意义, 其中有三个与J结合在一起的张量不变量:

$$J_1 = \mathrm{tr} J = 2\sum_k m_k(x_k^2 + y_k^2 + z_k^2),$$

$$J_2 = \begin{vmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{vmatrix} + \begin{vmatrix} J_{11} & J_{13} \\ J_{31} & J_{33} \end{vmatrix} + \begin{vmatrix} J_{22} & J_{23} \\ J_{32} & J_{33} \end{vmatrix},$$

$$J_3 = \det J.$$

J_1, J_2, J_3相对于旋转的不变性是J的特征多项式的不变性的一个推论.

把J化到主轴上去, J的对称性允许这样做, 我们就得到一个特征值$\lambda_i > 0$, i=1, 2, 3的矩阵$\mathrm{diag}\{\lambda_1, \lambda_2, \lambda_3\}$, 这些特征值被称为**惯性主矩**. 特别地, 惯性张量是正定的. 如果ω是刚体旋转的角速度, 而$\mathbf{j}$是角矩, 那么$\mathbf{j} = J\omega$, $\mathbf{j}$与ω成比例的性质成立, 当且仅当, 刚体围绕自己的一个主轴旋转.

在第1章§4的定理3, 当时, 我们在二次型和对称的双线性型之间建立了一个双射对应. 这个对应的较弱的形式就是在多线性型的情形也仍然能继续存在.

定义4 称V上的函数$Q : \mathbf{x} \mapsto \mathcal{K}$是$p$**次齐次函数**, 如果

$$Q(\mathbf{x}) = F(\mathbf{x}, \cdots, \mathbf{x}),$$

其中$F : V^p \to \mathcal{K}$是V上的任意一个p线性型.

把对称化映射应用到p线性型F上, 按定理2, 我们就得到一个对称的p线性型$S(F)$:

$$(S(F))(\mathbf{x}_1, \cdots, \mathbf{x}_p) = \frac{1}{p!} \sum_{\pi \in S_p} F(\mathbf{x}_{\pi 1}, \ldots, \mathbf{x}_{\pi p}),$$

其中$F: V^p \to \mathfrak{K}$是V上某一个p线性型.

明显地,

$$Q(\mathbf{x}) = (S(F))(\mathbf{x}, \cdots, \mathbf{x}). \tag{13}$$

称型$S(F)$是**与Q配极的p线性型**.

我们就部分地得到了下面的命题.

定理3 每个次数为p的齐次函数都可以经自己的配极p线性型表示成(13)形式. 与Q配极的线性型是唯一的.

证明 配极型的唯一性可由自己的坐标表示得到, 设

$$(S(F))(\mathbf{x}_1, \cdots, \mathbf{x}_p) = \sum (S(F))_{i_1 \cdots i_p} x_1^{i_1} \cdots x_p^{i_p}.$$

那么,

$$Q(\mathbf{x}) = \sum (S(F))_{i_1 \cdots i_p} x^{i_1} \cdots x^{i_p}. \tag{14}$$

n个未知量的p次齐次多项式$f(X_1, \cdots, X_n)$当$X_i = x^i$时的值是$Q(\mathbf{x})$, 而且有唯一的方法记成

$$f(X_1, \cdots, X_n) = \sum_{i_1 \leqslant \cdots \leqslant i_p} f_{i_1 \cdots i_p} X_{i_1} \ldots X_{i_p}. \tag{15}$$

比较(14)和(15), 我们得到

$$f_{i_1 \cdots i_p} = c \cdot (S(F))_{i_1 \cdots i_p},$$

其中$c = c(i_1, \cdots, i_p) \in \mathbb{Z}$是取所有的指标$i_1, \cdots, i_p$的置换一个不漏地产生的排列的个数. 例如, 当$p$=4时, $c(i, j, k, l) = 12$, $c(i, i, i, j) = 4$, 等等. $f_{i_1, \cdots, i_p}$被唯一确定, 随后, 系数$S(F)_{i_1 \cdots i_p}$亦被唯一确定. □

实际上, 我们得到了对称张量空间$\mathbb{T}_p^+(V)$与n个变元的p次齐次多项式(型)空间 $\mathfrak{K}[X_1, \cdots, X_n]_p$之间的一个双射对应. 对于空间$\mathbb{T}_+^p(V)$也有同样的事实. 我们在这种联系中应该注意到

$$\dim \mathfrak{K}[X_1, \cdots, X_n]_p = \binom{n+p-1}{p}.$$

4. 斜对称张量 和前面一样, 我们只限于讨论$(p, 0)$型和$(0, q)$型张量, 且把对称群S_p利用公式(10)或者(11)作用到$T \in \mathbb{T}_p^0(V)$.

定义5 称张量T是**斜对称的**或者**反对称的**, 如果

$$f_\pi(T) = \varepsilon_\pi T, \forall \pi \in S_p, \tag{16}$$

其中ε_π是置换π的(奇偶性)符号.

我们提醒, $\varepsilon : \pi \mapsto \varepsilon_\pi$是群$S_p$到$\{\pm 1\}$的一个同态且在所有的对换$\tau$上的值$\varepsilon_\tau = -1$.

条件(16)可以用等价要求

$$f_\tau(T) = -T \tag{16'}$$

来代替, 其中τ取遍对换的集合(由定义和等式$f_\sigma(f_\tau(T)) = f_{\sigma\pi}(T)$显见可以表达成

$$T(\cdots, \mathbf{x}, \cdots, \mathbf{y}, \cdots) = -T(\cdots, \mathbf{y}, \cdots, \mathbf{x}, \cdots), \quad \mathbf{x}, \mathbf{y} \in V \tag{17}$$

的形式. 带省略号的地方可以是任意向量, 但在(17)的两侧是一致的, 在我们这里有char $\mathfrak{K} = 0$, 所以, 当$\mathbf{x} = \mathbf{y} = \mathbf{z}$时, 由(17)可以得到

$$T(\cdots, \mathbf{z}, \cdots, \mathbf{z}, \cdots) = 0. \tag{17'}$$

在(17′)中令$\mathbf{z} = \mathbf{x} + \mathbf{y}$并利用$T$的多重线性性质, 可得

$$\begin{aligned} &T(\cdots, \mathbf{x}+\mathbf{y}, \cdots, \mathbf{x}+\mathbf{y}, \cdots) \\ =&T(\cdots, \mathbf{x}, \cdots, \mathbf{x}, \cdots) + T(\cdots, \mathbf{y}, \cdots, \mathbf{y}, \cdots) + T(\cdots, \mathbf{x}, \cdots, \mathbf{y}, \cdots) + \\ &T(\cdots, \mathbf{y}, \cdots, \mathbf{x}, \cdots), \end{aligned}$$

而这意味着, 由(17′)可以推出(17). 所以, (17)和(17′)是等价的.

张量T的斜对称性可以用坐标$T_{i_1\cdots i_p}$语言的明显的方式表达出来. 例如, 当p=2时, 斜对称就意味着$T_{ij} = -T_{ji}$ (线性算子或者n阶方阵的反对称性), 同时, 这个性质正是我们期待得到的, 它与基底的选择没有关系. 在一般情形

$$T_{i_{\pi 1}\cdots i_{\pi p}} = \varepsilon_\pi T_{i_1\cdots i_p}.$$

这样一来, 反对称张量的任意坐标就唯一确定了, 只要确定了一个排列下的坐标即可, 比方说, 按递增顺序的排列

$$T_{i_1 i_2 \cdots i_p}, 1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n. \tag{18}$$

所有的带有两个相同指标的坐标都等于零. 恰恰相反, 形如(18)的不同的配套的指标之间没有任何关系, 所以总共有$\binom{n}{p}$个无关系的指标. 我们立刻就可以把原先给的表达式弄得更精致一些.

定义6 称映射

$$A = \frac{1}{p!} \sum_{\pi \in S_p} \varepsilon_\pi f_\pi : \mathbb{T}_p^0(V) \to \mathbb{T}_p^0(V) \tag{19}$$

是**交错化**映射.

同样可以自然地引进斜对称张量的集合$\Lambda^p(V^*)$, 它包含在$\mathbb{T}_p^0(V)$中(相应地, $\Lambda^q(V) \subset \mathbb{T}_0^q$). 事实上, 这个集合还是个子空间, 比方说, 它可以下面的蕴涵式

$$f_\pi P = \varepsilon_\pi P, f_\pi R = \varepsilon_\pi R \Rightarrow f_\pi(\alpha P + \beta R), = \alpha f_\pi P + \beta f_\pi R \\ = \alpha\varepsilon_\pi P + \beta\varepsilon_\pi R = \varepsilon_\pi(\alpha P + \beta R), \ \ \forall \alpha, \beta \in \mathfrak{K},$$

看出来.

定理4 交错化映射(19)是个线性算子,

$$A(\alpha T + \beta R) = \alpha A(T) + \beta A(R),$$

且具有下列性质:

i) $A^2 = A$;

ii) $\operatorname{Im} A = \Lambda^p(V^*)$;

iii) $A(f_\sigma(T)) = \varepsilon_\sigma A(T)$.

证明 i) 按照(19)我们有

$$A^2 = \frac{1}{(p!)^2} \sum_{\sigma,\pi \in S_p} \varepsilon_\sigma \varepsilon_\pi f_\sigma \circ f_\pi = \frac{1}{(p!)^2} \sum_{\sigma,\pi \in S_p} \varepsilon_{\sigma\pi} f_{\sigma\pi} = \frac{1}{p!} \sum_{\rho \in S_p} \varepsilon_\rho f_\rho = A.$$

这里用到了一个事实, 任意元素$\rho \in S_p$都有$p!$种方式被表达成$\sigma\pi$的样子:σ是任意选的, 再找出一个π使得$\pi = \sigma^{-1}\rho$. 同样也用到了ε_σ和f_σ按照σ的可乘性质.

ii) 对所有的$T \in \mathbb{T}_p^0$, 有

$$f_\sigma(A(T)) = \frac{1}{p!} \sum_{\pi \in S_p} \varepsilon_\pi f_\sigma(f_\pi(T)) = \varepsilon_\sigma \frac{1}{p!} \sum_{\pi \in S_p} \varepsilon_{\sigma\pi} f_{\sigma\pi}(T) = \varepsilon_\sigma A(T),$$

所以, $\operatorname{Im} A \subset \Lambda_p(V^*)$. 另一方面,

$$T \in \Lambda^p(V^*) \Rightarrow A(T) = \frac{1}{p!} \sum_{\pi \in S_p} \varepsilon_\pi f_\pi(T) = \frac{1}{p!} \sum_{\pi \in S_p} \varepsilon_\pi^2 T = \frac{1}{p!} \sum_{\pi \in S_p} T = T.$$

这就给出了所需要的论断.

iii) 在ii)里已经验证了, $f_\sigma A = \varepsilon_\sigma A$, 按同样的道理, $A f_\sigma = \varepsilon_\sigma A$.

5. 张量空间 下面我们要贴近在不同的数学分支中被广泛使用的术语.

定义7 称共变的斜对称张量, 也就是$\Lambda^p(V^*)$的元素是一个**外p形式**, 或者, 更方便些, 称为V上的**p次外形式**. 反变的斜对称张量($\Lambda^p(V)$的元素)被称为**p向量**.

默认, $\Lambda'(V^*) = V^*$和$\Lambda'(V) = V$. 然后, 我们引入无穷多个张量空间$\mathbb{T}_p^0(V), p = 0, 1, \cdots$ 的外直和(见第1章§2):

$$\mathbb{T}(V^*) = \mathfrak{K} \oplus \mathbb{T}_1^0(V) \oplus \mathbb{T}_2^0(V) \oplus \cdots \tag{20}$$

这个和的元素可认为是序列

$$(f_0, f_1, f_2, \cdots) = \sum_{i\geqslant 0} f_i, \quad f_i \in \mathbb{T}_i^0(V),$$

它的项几乎全为零(即除有限处以外, 其余的都为零). 可以把$\mathbb{T}(V^*)$看成是对张量定义了乘法

$$\left(\sum_{i=0}^{s} f_i\right) \otimes \left(\sum_{j=0}^{t} g_j\right) = \sum_{i+j=0}^{s+t} f_i \otimes g_j = \sum_{k=0}^{s+t} h_k \tag{21}$$

的一个无穷维的结合代数. 这时, 自然地, 满足条件

$$\lambda(f \otimes g) = \lambda f \otimes g = f \otimes \lambda g, \quad \lambda \in \mathfrak{K}.$$

我们称$\mathbb{T}(V^*)$是**共变张量代数**.

规则(21)令人完全想起了多项式的乘法规则, 与其不同的仅仅是非交换性. 在 $\mathbb{T}_p^0(V)$中派生出对称张量子空间$\mathbb{T}_p^+(V)$, 正如我们在第3目末尾看到的, 它可以和n个变元的p次齐次多项式的空间等同起来. 它们的外直和就是通常的多项式代数, 而用适当方式再次定义的规则(21)与多项式的乘法是一致的.

用完全类似的方法引进**反变张量代数**

$$\mathbb{T}(V) = \mathfrak{K} \oplus \mathbb{T}_0^1 \oplus \mathbb{T}_0^2 \oplus \cdots, \tag{22}$$

而在它的里面对派生出一个对称张量的子空间

$$S(V) = T_+(V) = \bigoplus_{q=0}^{\infty} \mathbb{T}_+^q(V).$$

同时, 它又同构于通常的n=dim V个变量的多项式代数. 按公式

$$T_1T_2 = S(T_1 \otimes T_2), \quad T_1 \in T_+^p(V), \quad T_2 \in T_+^q(V).$$

在空间$S(V)$上引入交换的结合代数的显形式.

域$\mathfrak{K}$上的代数$S(V)$被称为空间V上的**对称代数**.

现在, 我们把注意力转向子空间

$$\Lambda(V^*) = \mathfrak{K} \oplus \Lambda^1(V^*) \oplus \Lambda^2(V^*) \oplus \cdots \subset \mathbb{T}(V^*).$$

$$\Lambda(V) = \mathfrak{K} \oplus \Lambda^1(V) \oplus \Lambda^2(V) \oplus \cdots \subset \mathbb{T}(V). \tag{23}$$

与共变的或反变的对称张量不同的是, 由(8)决定的子空间并不是$\mathbb{T}(V^*)$和相应的$\mathbb{T}(V)$的子代数. 但是, 这也说明, 在$\Lambda(V^*)$和$\Lambda(V)$中可以用自然的方式引进一个运算, 使得它们构成一个结合代数.

说明 我们研究过的对称化和交错化运算只能发生在特征数为零的域上. 在[2]中已经指明了如何摆脱这个限制.

习 题

1. 指标的提升与下放 照旧, 设(V, g)是个欧几里得空间, T是个(p,q)型张量, 坐标是$T^{j_1 j_2\cdots j_q}_{i_1 i_2\cdots i_p}$. 用坐标建立一个张量$\sum\limits_k g^{ik} T^{j_1 j_2\cdots j_q}_{k\, i_2\cdots i_p}$, 而且在得到的张量上用$i_1$代表指标$i$, 然后, 令

$$T^{i_1 j_1 j_2\cdots j_q}_{i_2\cdots i_p} = \sum_k g^{i_1 k} T^{j_1 j_2\cdots j_q}_{k i_2\cdots j_p}.$$

人们说, 在这个等式左侧的张量是**提升指标**i_1的算子作用在T上的结果. 借助于度量张量G的定义, **下放指标**算子可类似地给出. 例如

$$T^{j_1 j_3\cdots j_q}_{i_1\cdots i_p j_2} = \sum_k g_{j_2 k} T^{j_1 k j_3\cdots j_q}_{i_1\cdots i_p}.$$

就是在张量T上把指标j_2下放到下方指标的最末位而得到的张量的坐标.

一般地, 提升第s个指标和下放第t个指标都是线性映射

$$\mathbb{T}^q_p(V) \to \mathbb{T}^{q+1}_{p-1} \quad \mathbb{T}^q_p \to \mathbb{T}^{q-1}_{p+1},$$

用了它们就得到以

$$\Gamma_{i_1\cdots i_{s-1}}{}^{i_s}{}_{i_{s+1}\cdots i_p}{}^{j_1\cdots j_q}, \quad T_{i_1\cdots i_p}{}^{j_1\cdots j_{t-1}}{}_{j_t}{}^{j_{t+1}\cdots j_q}$$

为系数的张量. 在这里, 在张量的坐标上我们使用了"分块排列"合成指标. 例如, $T\in V\otimes V^*$, 可以给出分量用T_i^j代表, 而$T\in V\otimes V^*\otimes V\otimes V$给出分量$T_{i_1}{}^{j_1}{}_{i_2 i_3}$.

指标提升算子和下放算子可以多次运用, 而且可以用到度量张量本身.

证明,

$$g^{ik} g^{jl} g_{kl} = g^{ij}, \quad g_{ik} g_{jl} g^{kl} = g_{ij}.$$

并验证等式$g^{ik} x_k = x^i$的正确性, 这意味着, 向量$\mathbf{x}$的反变坐标x^i可以把在同一个向量的共变指标x_i上施以下放指标算子得到.

2. 设$\dim V > 1$. 证明, $\mathbb{T}(V)$是一个无零因子的交换环. 在$\mathbb{T}(V)$中非零的纯量是仅有的可逆元.

3. 在例2中计算出张量不变量J_2和J_3的显式.

§3 外代数

1. 外积 就像前面的叙述方式一样, 可以毫无差别地, 可以用外形式和p向量语言来叙述. 为了多样性, 取空间$\Lambda(V)$.

定义1 对任意q向量Q和任意r向量R, 令

$$Q\wedge R = A(Q\otimes R), \tag{1}$$

就给定了**外积运算**

$$\wedge : \Lambda(V)\times\Lambda(V)\to\Lambda(V).$$

在这里, 把$A(Q\otimes R)$理解为把交错化算子作用到张量$Q\otimes R$所得的结果. 因为$Q\otimes R\in\mathbb{T}_0^{q+r}(V)$, 而张量$A(Q\otimes R)$按照§4的定理4是斜对称的, 所以, 公式(1)给出了映射

$$\wedge:\Lambda^q(V)\times\Lambda^r(V)\to\Lambda^{q+r}(V)$$

(默认, $\lambda\wedge R=R\wedge\lambda=\lambda R,\forall\lambda\in\mathfrak{K}$). 为了给出空间$\Lambda(V)$中任意两个元素$Q'$和$R'$的外积$Q'\wedge R'$, 我们记

$$Q'=\sum_{i\geqslant 0}Q_i,\quad R'=\sum_{j\geqslant 0}R_j;\quad Q_i\in\Lambda^i(V),\quad R_j\in\Lambda^j(V),$$

并且令

$$Q'\wedge R'=\sum_{i+j\geqslant 0}Q_i\wedge R_j.$$

还要注意, 如果在(1)中用$\alpha R+\beta T\in\Lambda^r(V)$代替$R$, 那么,

$$\begin{aligned}Q\wedge(\alpha R+\beta T)&=A(Q\otimes(\alpha R+\beta T))=A(\alpha Q\otimes R+\beta Q\otimes T)\\&=\alpha A(Q\otimes R)+\beta A(Q\otimes T)=\alpha(Q\wedge R)+\beta(Q\wedge T).\end{aligned}$$

类似地,

$$(\alpha Q+\beta S)\wedge R=\alpha(Q\wedge R)+\beta(S\wedge R).$$

2. 向量空间的外代数 我们已经建立了一个具有双线性性质(或者, 可分配的)乘法运算, 也就是说, 这种运算在$\Lambda(V)$给出了一个代数结构.

定义2 称域$\mathfrak{K}$上的代数$\Lambda(V)$是空间V的**外代数**(或者**格拉斯曼代数**).

代数$\Lambda(V)$有单位元素, 它等同于$1\in\mathfrak{K}$. 外代数的更重要的性质的结论将在下面的定理2中给出.

定理1 对任意张量$Q\in\mathbb{T}_0^q(V)$和$R\in\mathbb{T}_0^r(V)$, 关系式

$$A(A(Q)\otimes R)=A(Q\otimes A(R))=A(Q\otimes R)$$

成立.

证明 注意, 按定义

$$A(Q)=\frac{1}{p!}\sum_{\pi\in S_p}\varepsilon_\pi f_\pi(Q),$$

进而,

$$A(A(Q)\otimes R)=\frac{1}{(q+r)!}\sum_{\sigma\in S_{q+r}}\varepsilon_\sigma f_\sigma(A(Q)\otimes R).$$

利用交错化算子的线性性, 我们得到

$$A(A(Q)\otimes R)=\frac{1}{q!}\sum_{\pi\in S_q}\varepsilon_\pi A(f_\pi(Q)\otimes R).\tag{2}$$

研究嵌入映射$\varphi : S_q \to S_{q+r}$, 并把$\tilde{\pi} = \varphi(\pi)$, $\pi \in S_p$理解为S_{q+r}的一个置换, 作用规则是

$$\tilde{\pi} i = \begin{cases} \pi i, & \text{如果} i \leqslant p, \\ i, & \text{如果} i > p. \end{cases}$$

那么, 就有$f_\pi(Q) \otimes R = f_{\tilde{\pi}}(Q \otimes R)$ (见§2的公式(11)), 而且按§2定理4的3), 有

$$A(f_\pi(Q) \otimes R) = Af_{\tilde{\pi}}(Q \otimes R) = \varepsilon_{\tilde{\pi}} A(Q \otimes R).$$

再注意到, $\varepsilon_{\tilde{\pi}} = \varepsilon_\pi$, 我们就可以把关系(2)化成所需要的形式

$$A(A(Q) \otimes R) = \frac{1}{q!} \sum_{\pi \in S_q} \varepsilon_\pi^2 A(Q \otimes R) = A(Q \otimes R),$$

因为$\varepsilon_\pi^2 = 1$. 类似地, 可以证明$A(Q \otimes A(R)) = A(Q \otimes R)$. □

定理2 外代数$\Lambda(V)$是结合的.

证明 我们只需验证恒等式

$$(P \wedge Q) \wedge R = P \wedge (Q \wedge R) \tag{3}$$

对任意$P, Q, R \in \Lambda(V)$. 由算子Λ有多重线性性质, 只需研究

$$P \in \Lambda^p(V), \quad Q \in \Lambda^q(V), \quad R \in \Lambda^r(V).$$

的情形, 就足够了. 按照(1),

$$(P \wedge Q) \wedge R = A(A(P \otimes Q) \otimes R),$$

而按定理1, 有

$$A(A(P \otimes Q) \otimes R) = A((P \otimes Q) \otimes R).$$

然后, 运用张量乘积的结合性, 并再次利用定理1, 就得到

$$\begin{aligned} A((P \otimes Q) \otimes R) &= A(P \otimes (Q \otimes R)) \\ &= A(P \otimes A(Q \otimes R)) = P \wedge (Q \wedge R). \end{aligned}$$

把所有这些都联系起来, 就给出了我们的恒等式(3). □

作为一个结合代数, 如同群一样, 无论如何配置括号都没有差别了, 所以, 乘积$P_{i_1} \wedge P_{i_2} \wedge \cdots \wedge P_{i_m}$有意义. 值得指出的是, 与(1)对应的

$$\mathbf{x} \wedge \mathbf{y} = \frac{1}{2}(\mathbf{x} \otimes \mathbf{y} - \mathbf{y} \otimes \mathbf{x}) = A(\mathbf{x} \otimes \mathbf{y}) \tag{4}$$

对任意$\mathbf{x}, \mathbf{y} \in V$都成立$(p = r = 1)$, 因此, 有

$$\mathbf{x} \wedge \mathbf{y} = -\mathbf{y} \wedge \mathbf{x}, \quad \mathbf{x} \wedge \mathbf{x} = 0. \tag{5}$$

推论 设$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$是$V$中的任意向量, 那么,

$$\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_p = A(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \cdots \otimes \mathbf{x}_p). \tag{6}$$

证明 当p=2时, 关系式(6)和(4)是一致的, 当$p > 2$时, 对p用归纳法. 在定理1和定理2的基础上, 我们得到

$$\begin{aligned}\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_p &= (\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_{p-1}) \wedge \mathbf{x}_p = A((\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_{p-1}) \otimes \mathbf{x}_p) \\ &= A(A(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_{p-1}) \otimes \mathbf{x}_p) = A(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_{p-1} \otimes \mathbf{x}_p).\end{aligned} \qquad \square$$

定理3 设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是向量空间$V$的基底, 那么, p向量

$$\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_p}, \ 1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n \tag{7}$$

组成空间$\Lambda^p(V)$的一个基底.

证明 任意一个乘积$\mathbf{e}_{j_1} \wedge \mathbf{e}_{j_2} \wedge \cdots \wedge \mathbf{e}_{j_p}$都可以把因子按序号大小规律排列起来, 或者, 当$j_r = j_s$时, 得到零(见(5)). 因而, 所有这种乘积都包含在形如(7)的张量的集合中了.

然后, 我们设$P \in \Lambda^p(V)$. 与$\mathbb{T}_0^p(V)$中的所有的张量一样, P可由基底表示出来

$$P = \sum_{j_1 \cdots j_p} P^{j_1 \cdots j_p} \mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p}.$$

由交错化映射的线性性, §2的定理4以及关系式(6), 我们有

$$P = A(P) = \sum_{j_1 \cdots j_p} P^{j_1 \cdots j_p} A(\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p}) = \sum_{j_1 \cdots j_p} P^{j_1 \cdots j_p} \mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_p},$$

而这意味着任意一个p向量都可以由形如(7)的向量表示出来, 剩下来需要我们证明它们的线性无关性. 设

$$\sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \lambda^{i_1 \cdots i_p} \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p} = 0.$$

那么, 由(6)得出

$$\sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \lambda^{i_1 \cdots i_p} A(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p}) = 0,$$

或者, 更方便地, 记成

$$\frac{1}{p!} \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \lambda^{i_1 \cdots i_p} \sum_{\pi \in S_p} \varepsilon_\pi f_\pi(\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p}) = 0.$$

利用§2的公式(11′), 我们得到

$$\sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \lambda^{i_1 \cdots i_p} \sum_{\pi \in S_p} \varepsilon_\pi (\mathbf{e}_{i_{\pi 1}} \otimes \cdots \otimes \mathbf{e}_{i_{\pi p}}) = 0. \tag{8}$$

如果$\pi \neq e$, 那么, 指标序列$i_{\pi 1}, \cdots, i_{\pi p}$就已经不是从小到大按顺序排列了. 可见, 关系式(8)可改写成

$$\sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \lambda^{i_1 \cdots i_p} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} + \cdots = 0 \tag{9}$$

的样子, 其中省略号代表指标序列$j_1, \cdots, j_p$有倒序的张量$\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_p}$的线性组合. 因为张量$\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p}$是线性无关的, 所以, 由(9)推出, $\lambda^{i_1 \cdots i_p} = 0$. □

回忆一下(见§2的(8)式), $\Lambda(V)$是子空间$\Lambda^p(V)$的外直和, 所以, 特别地, 有

$$\dim \Lambda(V) = \sum_{p \geqslant 0} \dim \Lambda^p(V),$$

进而有

推论 空间V的外代数$\Lambda(V)$的维数是2^n, 同时

$$\dim \Lambda^p(V) = \binom{n}{p}.$$

空间$\Lambda^n(V)$的基底由一个n向量

$$\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_n$$

生成.

证明 足够明显, 因为p向量$\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}$的个数$i_1 < \cdots < i_p$时等于$n$中取$p$个的排列的个数. 其次, 当$p > n$时, $\binom{n}{p} = 0, \binom{n}{n} = 1$, 且

$$\sum_{p=0}^{n} \binom{n}{p} = 2^n.$$

□

我们还需要注意外乘积算子的一个有用的性质, 即

$$Q \in \Lambda^q(V),\ R \in \Lambda^r(V) \Rightarrow Q \wedge R = (-1)^{qr} R \wedge Q. \tag{10}$$

鉴于关系式(10)的形式, 通常说, 空间V上的外代数是(普通的)**反交换的**.

由于算子Λ具有双线性性质, 要证明(10)式, 只要考察情形

$$Q = \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_q}, \quad R = \mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_r}.$$

就足够了. 把关系式(5)利用q次, 我们就得到

$$(\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_q}) \wedge \mathbf{e}_{j_k} = (-1)^q \mathbf{e}_{j_k} \wedge (\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_q}),$$

也就是$Q \wedge \mathbf{e}_{j_k} = (-1)^q \mathbf{e}_{j_k} \wedge Q$与$j_k$是否和某一个$i_s$重合没有关系. 这样一来.

$$\begin{aligned} Q \wedge R &= Q \wedge (\mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_r}) = (-1)^q \mathbf{e}_{j_1} \wedge (Q \wedge \mathbf{e}_{j_2} \wedge \cdots \wedge \mathbf{e}_{j_r}) \\ &= (-1)^{2q} \mathbf{e}_{j_1} \wedge \mathbf{e}_{j_2} \wedge (Q \wedge \mathbf{e}_{j_3} \wedge \cdots \wedge \mathbf{e}_{j_r}) = \cdots = (-1)^{qr} R \wedge Q. \end{aligned}$$

由关系式(10), 当p为奇数时, 对任一个p向量, 可以得到

$$P \wedge P = 0. \tag{11}$$

在p为偶数时, 关系式(11)就不一定能被满足.

例1 设$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_n \rangle, n \geqslant 4$. 那么

$$\begin{aligned}&(\mathbf{e}_1 \wedge \mathbf{e}_2 + \mathbf{e}_3 \wedge \mathbf{e}_4) \wedge (\mathbf{e}_1 \wedge \mathbf{e}_2 + \mathbf{e}_3 \wedge \mathbf{e}_4)\\ = &\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge (\mathbf{e}_3 \wedge \mathbf{e}_4) + \mathbf{e}_3 \wedge \mathbf{e}_4 \wedge (\mathbf{e}_1 \wedge \mathbf{e}_2) = 2\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3 \wedge \mathbf{e}_4 \neq 0.\end{aligned}$$

3. 与行列式的联系 利用外积运算很容易刻画向量组线性无关的判别方法和重新得到行列式的所有重要性质.

定理4 *设V是域$\mathfrak{K}$上的一个n维向量空间, 为使任意p个向量$\mathbf{x}_1, \cdots, \mathbf{x}_p \in V$线性无关, 充要条件是*

$$\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_p \neq 0.$$

证明 如果向量$\mathbf{x}_1, \cdots, \mathbf{x}_p$构成一个线性相关组, 那么, 至少有一个, 比方说是$\mathbf{x}_p$, 是其余向量的线性组合. 这个时候, 外积$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p$可以分解成外积

$$\alpha_i \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_{p-1} \wedge \mathbf{x}_i,\ i < p$$

的和, 而其中每一项都含有两个相同的因子, 从而必等于零.

相反, 如果$\mathbf{x}_1, \cdots, \mathbf{x}_p$ 线性无关, 那么, 我们可以把它们取作某一个基底的前边p 个向量. 于是, p 向量$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p$ 相应地就是$\Lambda^p(V)$ 的基底中的一个元素, 从而它不为零. □

设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ 是空间V 的一个基底, 而$\mathbf{x}_1, \cdots, \mathbf{x}_n$ 是V 中的任意一组向量. 我们有

$$\mathbf{x}_j = \sum_{i=1}^{n} x_j^i \mathbf{e}_i, \quad j = 1, \cdots, n.$$

按照定理3的推论

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n = \Delta \cdot \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n,$$

其中$\Delta = \Delta(x_j^i)$是向量$\mathbf{x}_1, \cdots, \mathbf{x}_n$的一个纯量函数, 或者, 同样地, 是矩阵$(x_j^i)$列的函数. 由外积的性质可以直接推出, 这个函数是线性的, 斜交换的, 而且对于$x_j = \mathbf{e}_j$, $j = 1, \cdots, n$有$\Delta = 1$. 行列式的唯一性定理[BA Ⅰ]第3章§1表明, $\Delta = \det(x_j^i)$. 这样一来, 就有

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n = \det(x_j^i) \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n. \tag{12}$$

在任意一个p维子空间$U \subset V$中, 对于它的任意两个基底$(\mathbf{a}_i)$, $(\mathbf{b}_i)$, 关系式(12)对应关系式

$$\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_p = \lambda \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_p, \quad \lambda \in \mathfrak{K}. \tag{12$'$}$$

说明 由关系式(12)可以毫无困难地推出行列式的所有性质. 更进一步, 这时可以获得一系列定理的最自然的证明, 外代数可以成为线性代数的有重要价值的分支的基础. 这个说明, 对我们来说, 是一个迟到的注解, 但是, 话又说回来了, 这一点并不值得婉惜.

对于一组p个向量$\mathbf{x}_j \in V$的类似的公式可以算是关系式(12)的一个推广, 即

$$\begin{aligned}\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p &= \sum_{i_i \cdots i_p} x_1^{i_1} \cdots x_p^{i_p} \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p} \\ &= \sum_{1 \leqslant i_i < \cdots < i_p \leqslant n} \Big(\sum_{\pi \in S_p} x_1^{i_{\pi 1}} \cdots x_p^{i_{\pi p}} \mathbf{e}_{i_{\pi 1}} \wedge \cdots \wedge \mathbf{e}_{i_{\pi p}}\Big) \\ &= \sum_{i_1 < \cdots < i_p} \Big(\sum_{\pi \in S_p} x_1^{i_{\pi 1}} \cdots x_p^{i_{\pi p}} (\varepsilon_\pi \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p})\Big) \\ &= \sum_{i_1 < \cdots < i_p} \Big(\sum_{\pi \in S_p} \varepsilon_\pi x_1^{i_{\pi 1}} \cdots x_p^{i_{\pi p}}\Big) \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}.\end{aligned}$$

按照[BA Ⅰ]第3章§1的公式(3), 我们得到行列式的完全展开式

$$\begin{aligned}\Delta_{i_1 \cdots i_p}(\mathbf{x}_1, \cdots, \mathbf{x}_p) := &\sum_{\pi \in S_p} \varepsilon_\pi x_1^{i_{\pi 1}} \cdots x_p^{i_{\pi p}} \\ = &\begin{vmatrix} x_1^{i_1} & \cdots & x_p^{i_p} \\ \cdots & \cdots & \cdots \\ x_1^{i_p} & \cdots & x_p^{i_p} \end{vmatrix} = \det(x_j^{i_k}).\end{aligned} \tag{13}$$

由此可见, 下面定理成立.

定理5 设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是空间$V$的一个基底, 而

$$\mathbf{x}_j = \sum_{i=1}^n x_j^i \mathbf{e}_i, \quad 1 \leqslant j \leqslant p,$$

是V中任意p个向量. 那么,

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p = \sum_{1 \leqslant i_i < \cdots < i_p \leqslant n} \Delta_{i_1 \cdots i_p}(\mathbf{x}_1, \cdots, \mathbf{x}_p) \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}, \tag{14}$$

其中$\Delta_{i_1 \cdots i_p}(\mathbf{x}_1, \cdots, \mathbf{x}_p)$是形如(13)的行列式.

当$p = n$时, 公式(14)就归结为(12). 注意, 在规范化条件$\Delta_{i_1 \cdots i_p}(\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_p}) = 1$之下, 可以把$T = \Delta_{i_1 \cdots i_p}$看成和在外$p$形式中是一样的. 特别地, $\det = \Delta_{1 \cdots n}$对于$V$上的外$n$形式的基底$(\mathbf{e}_i)$是唯一的, 这个时候, $\det(\mathbf{e}_1, \cdots, \mathbf{e}_n) = 1$.

4. 向量子空间与p向量 照例, 设V是域$\mathcal{K}$上的一个n维向量空间.

定义3 称$\Lambda^p(V)$中的$P \neq 0$决定的子空间

$$\operatorname{Ann} P = \{\mathbf{x} \in V | P \wedge \mathbf{x} = \mathbf{0}\}$$

是p向量P的**零化子**, 还约定, 称p向量P是**可分解的**, 如果有V中的向量$\mathbf{a}_i \in V$, $i = 1, 2, \cdots, p$使得

$$P = \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \cdots \wedge \mathbf{a}_p.$$

事实上, 集合Ann P是个向量子空间, 这很容易直接验证

$$\begin{aligned}\mathbf{x}_1, \mathbf{x}_2 \in \operatorname{Ann} P \Rightarrow P \wedge (\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) &= \alpha_1 P \wedge \mathbf{x}_1 + \alpha_2 P \wedge \mathbf{x}_2 = 0 \\ \Rightarrow \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 &\in \operatorname{Ann} P, \quad \forall \alpha_1, \alpha_2 \in \mathfrak{K}.\end{aligned}$$

定理6　设Ann $P = \langle \mathbf{e}_1, \cdots, \mathbf{e}_r \rangle$是个$r$维子空间. 那么, $r \leqslant P$且存在一个$(P - r)$向量Q使得

$$P = \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_r \wedge Q.$$

等式$r = P$ 成立, 当且仅当, p向量P是可分解的.

证明　首先, 我们不妨假设

$$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_r; \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle.$$

记

$$P = \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} P^{i_1 \cdots i_p} \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{i_p}, \tag{15}$$

我们把零化子条件表达成显形式

$$\sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} P^{i_1 \cdots i_p} \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_j = 0, \quad j = 1, 2, \cdots, r. \tag{16}$$

如果j与某个i_k一致, 那么, $\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_j = 0$, 因而得(16)的左侧只剩下$i_1 \neq j, \cdots, i_p \neq j$的被加项. 因为$(p+1)$向量$\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p} \wedge \mathbf{e}_j$指标不相同, 线性无关, 所以, 公式(16)中, 只要这组指标不整个地包含$1, 2, \cdots, r$, 就必然有$P^{i_1 \cdots i_p} = 0$. 按条件$P \neq 0$, 所以$r \leqslant p$而且

$$\begin{aligned} P &= \sum P^{12 \cdots r i_{r+1} \cdots i_p} \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_r \wedge \mathbf{e}_{i_{r+1}} \wedge \cdots \wedge \mathbf{e}_{i_p} \\ &= \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_r \wedge Q, \end{aligned} \tag{17}$$

其中

$$Q = \sum P^{12 \cdots r i_{r+1} \cdots i_p} \mathbf{e}_{r+1} \wedge \cdots \wedge \mathbf{e}_{i_p}$$

是某个$(p - r)$向量.

现在, 如果$r = p$, 那么, 对P可得到一个表达式$P = \lambda \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_p$, 它意味着, P是个可分解的p向量.

反过来, 若$P \neq 0$是可分解的, 即

$$P = \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \cdots \wedge \mathbf{a}_p,$$

那么, $\mathbf{a}_i \in \text{Ann } P$, $i = 1, 2, \cdots, p$. 也就是, $\dim \text{ Ann } P \geqslant p$, 而又因为在任何时候都有$\dim \text{ Ann } P \leqslant p$, 所以, $\dim \text{ Ann } P = p$. □

定理7 设$U = \langle \mathbf{a}_1, \cdots, \mathbf{a}_p \rangle$, $W = \langle \mathbf{b}_1, \cdots, \mathbf{b}_p \rangle$是$V$的具有相同维数$p$的子空间. 那么, U和W重合的充要条件是p向量$P = \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_p$和$p$向量$Q = \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_p$是成比例的$p$向量.

证明 正如我们已经看到的(定理6), $U = \text{Ann } P = \langle \mathbf{a}_1, \cdots, \mathbf{a}_p \rangle$. 如果同样$\langle \mathbf{b}_1, \cdots, \mathbf{b}_p \rangle = \text{Ann } P$, 那么, 由(12′)式就有$P = \lambda Q$. 反过来:如果$P = \lambda Q$, 那么

$$U = \text{Ann } P = \text{Ann } Q = W. \qquad \square$$

下面的定理对定理6和定理7做更详细的说明.

定理8 设$U = \text{Ann } P$, $W = \text{Ann } Q$, 其中P是可分解的p向量而Q是可分解的q向量, 那么

i) $U \supseteq W \Leftrightarrow P = Q \wedge R$ (R是某个$(p-q)$向量);

ii) $U \cap W = 0 \Leftrightarrow P \wedge Q \neq 0$.

当$U \cap W = 0$时, $U \oplus W = \text{Ann}(P \wedge Q)$.

证明 断言i)已经包含在定理6中了. 其次, 如果$P = \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_p$, $Q = \mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_q$且$P \wedge Q \neq 0$, 那么, 向量$\mathbf{a}_i$, $\mathbf{b}_j$, $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant q$是线性无关的且构成空间$U \oplus W$的一个基底. 如果$P \wedge Q = 0$, 那么就存在非平凡的线性关系式

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_p \mathbf{a}_p + \beta_1 \mathbf{b}_1 + \cdots + \beta_q \mathbf{b}_q = 0,$$

由此可推出$U \cap W \neq 0$.

5. p向量可分解条件 定理6—定理8指明了p向量的特殊作用. 在别的问题中也会遇到它们. p向量$P \in \Lambda^p(V)$的可分解性的条件可以用它的分量$P^{i_1 \cdots i_p}$的代数关系式加以刻画(见(15)). 事实上, 设$\mathbf{x} = \sum\limits_i x^i \mathbf{e}_i$是任意一个向量, 乘积$P \wedge \mathbf{x}$就是以$x^1, \cdots, x^n$的线性型为坐标的$(p+1)$向量. $(p+1)$向量的分量个数是$\begin{pmatrix} n \\ p+1 \end{pmatrix}$. 因此, 条件$\mathbf{x} \in \text{Ann } P$等价于满足一个$n$个不定元的$\begin{pmatrix} n \\ p+1 \end{pmatrix}$个方程的齐次方程组. 这个方程组的解空间与$\text{Ann } P$是一致的. 按定理6, 维数$r \leqslant p$, 并且, $r = p$的充要条件是P为可分解的. 换句话说, P可分解的充要条件是方程组的矩阵的秩不超过$n - p$, 也就是所有的$n - p + 1$阶的子式都为零. 这样就有了下面的可分解性的条件.

定理9 所有的$(n-1)$向量$P \neq 0$都是可分解的.

证明 根据(12)

$$P \wedge \mathbf{x} = f(\mathbf{x}) \cdot \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n, \quad f(\mathbf{x}) \in \mathfrak{K},$$

其中f是个纯量函数, 而$(\mathbf{e}_i)$是V的一个基底. 由外积运算的性质可以推出f是个线性函数. 此外, $\text{Ann } P = \text{Ker } f$, 从而就有$\dim \text{ Ann } P = \dim \text{ Ker } f \geqslant n - 1$. 按定

理6, 我们又有$\dim\ \text{Ann}\ P \leqslant n-1$, 因此$f \neq 0$. 进而$\dim\ \text{Ker}\ f \leqslant n-1$, 可见, $\dim\ \text{Ann}\ P = n-1$. □

例2 设V是个3维的欧几里得空间, ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$)是它的一个标准正交基底. 如果$P = \mathbf{a} \wedge \mathbf{b} \neq 0$, 那么,

$$\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{x} = (\mathbf{c}|\mathbf{x}) \cdot \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3,$$

其中$(\mathbf{c}|\mathbf{x}) = f(\mathbf{x})$. 可以验证$\mathbf{c} = [\mathbf{a}, \mathbf{b}]$是向量$\mathbf{a}$, $\mathbf{b}$向量积.

显然, 对每个可分解的p向量P都有$P \wedge P = 0$. 还可以指出, 对于二重向量的情形($p = 2$)条件$P \wedge P = 0$还是可分解的充分条件.

定理10 二重向量

$$P = \sum_{1 \leqslant i < j \leqslant n} P^{ij} \mathbf{e}_i \wedge \mathbf{e}_j \tag{18}$$

是可分解的, 当且仅当, $P \wedge P = 0$.

证明 需要证明, 如果$P \wedge P = 0$, 那么P必然是可分解的. 对$n = \dim\ V$用归纳法. 当$n = 3$时, 命题可由定理9推出来.

把(18)中所有包含$\mathbf{e}_1$的项放到一起, 它们可以表达成$\mathbf{e}_1 \wedge \mathbf{a}$的形式, 再设

$$Q = \sum_{2 \leqslant i < j \leqslant n} P^{ij} \mathbf{e}_i \wedge \mathbf{e}_j, \tag{19}$$

我们就有

$$P = \mathbf{e}_1 \wedge \mathbf{a} + Q,$$

而且可以认为$\mathbf{a}$是$\mathbf{e}_2, \cdots \mathbf{e}_n$的一个线性组合, 而$Q \neq 0$. 由$P \wedge P = 0$可以推知

$$(\mathbf{e}_1 \wedge \mathbf{a}) \wedge Q + Q \wedge (\mathbf{e}_1 \wedge \mathbf{a}) + Q \wedge Q = 0.$$

再借助于关系式(10), 我们由此得出

$$2\mathbf{e}_1 \wedge \mathbf{a} \wedge Q + Q \wedge Q = 0. \tag{20}$$

等式(20)左端的所有4向量都可由基底向量$\mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k \wedge \mathbf{e}_l$表示出来. 在$\mathbf{e}_1 \wedge \mathbf{a} \wedge Q$中包含因子$\mathbf{e}_1$, 而如同在(19)看到的, 在$Q$中只包含项$\mathbf{e}_i \wedge \mathbf{e}_j$, $i, j \geqslant 2$, 所以, 在$Q \wedge Q$中不会遇到因子$\mathbf{e}_1$. 因而, 等式(20)可以分成两个等式

$$\mathbf{e}_1 \wedge \mathbf{a} \wedge Q = 0, \quad Q \wedge Q = 0. \tag{21}$$

但是, $Q \in \Lambda^2(U)$, 其中$U = \langle \mathbf{e}_2, \cdots, \mathbf{e}_n \rangle$. 因此, 对$Q$可使用归纳法的假定, 即它是可分解:$Q = \mathbf{b} \wedge \mathbf{c}$. 把这种表达代入到(21)的第一个等式中, 我们得到

$$\mathbf{e}_1 \wedge \mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} = 0. \tag{22}$$

由这个乘积的四个因子的线性相关性, 只能有$\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}\in U$. 如果$\mathbf{b}=\lambda\mathbf{a}$, 那么, $P=\mathbf{e}_1\wedge\mathbf{a}+\lambda\mathbf{a}\wedge\mathbf{c}=(\mathbf{e}_1-\lambda\mathbf{c})\wedge\mathbf{a}$, 证明就完成了. 在此基础上, 我们不妨设$\mathbf{e}_1$, $\mathbf{a}$, $\mathbf{b}$是线性无关的, 按定理4, 我们有$R=\mathbf{e}_1\wedge\mathbf{a}\wedge\mathbf{b}\neq 0$, 再依据定理3, 在这个情形, $\dim\,\mathrm{Ann}\,R=3$. 更精确地, $\mathrm{Ann}\,R=\langle\mathbf{e}_1,\mathbf{a},\mathbf{b}\rangle$, 但是与(22)相对应, $\mathbf{c}\in\mathrm{Ann}\,R$, 这意味着, $\mathbf{c}=\alpha\mathbf{a}+\beta\mathbf{b}$. 在这种情况下

$$Q=\mathbf{b}\wedge\mathbf{c}=\mathbf{b}\wedge(\alpha\mathbf{a}+\beta\mathbf{b})=\alpha\mathbf{b}\wedge\mathbf{a}$$

且

$$P=\mathbf{e}_1\wedge\mathbf{a}+Q=\mathbf{e}_1\wedge\mathbf{a}+\alpha\mathbf{b}\wedge\mathbf{a}=(\mathbf{e}_1+\alpha\mathbf{b})\wedge\mathbf{a}.$$ □

推论 二重向量

$$P=P^{12}\mathbf{e}_1\wedge\mathbf{e}_2+P^{13}\mathbf{e}_1\wedge\mathbf{e}_3+P^{14}\mathbf{e}_1\wedge\mathbf{e}_4+P^{23}\mathbf{e}_2\wedge\mathbf{e}_3+P^{24}\mathbf{e}_2\wedge\mathbf{e}_4+P^{34}\mathbf{e}_3\wedge\mathbf{e}_4$$

是可分解的, 当且仅当, 它的坐标P^{ij}能用关系式

$$P^{12}P^{34}-P^{13}P^{24}+P^{14}P^{23}=0 \tag{23}$$

联系起来.

现在, 设$\mathbb{P}^3=\mathbb{P}(V)$是由4维实向量空间$V$生成的3维射影空间. 我们记得, 在$\mathbb{P}^3$中的直线对应$V$的2维平面. 按定理6, 任意平面$U<V$都是一个可分解的二重向量$P$的零化子, P是完全确定的(不计纯量). 换句话说, 二重向量P的坐标$(P^{12}:P^{13}:\cdots:P^{34})$可以看成是$\mathbb{P}^3$中直线的齐次坐标(**普吕克直线坐标**). 这些坐标, 作为6元组同样决定了5维空间$\mathbb{RP}^5$的一个点. 我们看到, 4维射影空间的直线相互单值地对应5维射影空间中由方程式(23)决定的一个点二次曲面.

习 题

1. 设$V=\mathbb{R}^n$是实空间, 由向量列$A^{(1)},\cdots,A^{(n)}$张成. 而B是任意一个向量列. 证明, 向量方程式

$$\sum_k\lambda_kA^{(k)}=B$$

的解参数λ_k可由关系式

$$(A^{(1)}\wedge\cdots\wedge A^{(n)})\lambda_k=A^{(1)}\wedge\cdots\wedge A^{(k-1)}\wedge B\wedge A^{(k+1)}\wedge\cdots\wedge A^{(n)}$$

给出.

试由此导出克拉默公式([BA I]第3章的§3).

2. 设V是特征数为0的域$\Bbbk$上的一个n维向量空间,

$$Z=\sum_{k=0}^{n}Z_k\ (Z_k\in\Lambda^k(V))$$

是外代数$\Lambda(V)$的一个元素,

$$\mathcal{Z}(\Lambda(V))=\{Z|Z\wedge X=X\wedge Z\quad\forall X\in\Lambda(V)\}$$

是这个外代数的中心.

证明

(1) Z是可逆的, 即Z^{-1}存在, 当且仅当, $Z_{(0)} \neq 0$;

(2) $Z \in \mathcal{Z}(\Lambda(V)) \Leftrightarrow \begin{cases} 如果n = 2m, 对任意i = 1, \cdots, m都有Z_{2i-1} = 0; \\ 如果n = 2m-1, 对任意i = 1, \cdots, m-1 都有 Z_{2i-1} = 0. \end{cases}$

在无穷维空间的情形如何从新给出上述条件?

3. 设$\mathcal{A} : V \to V$是个线性算子. 线性算子

$$\wedge^p \mathcal{A} : \Lambda^p(V) \to \Lambda^p(V)$$

由可分解的p向量定义

$$(\wedge^p \mathcal{A})(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p) = \mathcal{A}(\mathbf{x}_1) \wedge \cdots \wedge \mathcal{A}(\mathbf{x}_p).$$

称$\Lambda^p \mathcal{A}$是$\mathcal{A}$的**外p次幂**. 它显然是个交错的多线性映射. 对于V的一个固定的基底就可以研究算子$\mathcal{A}$的矩阵A和这个矩阵的外p次幂$\Lambda^p A$. 证明

$$\det\ \Lambda^p A \cdot \det\ \Lambda^{n-p} A = (\det\ A)^{\binom{n}{p}}$$

4. 设在向量空间V上给定了一个非退化的二次型q. 再设f是与q配极的双线性型. 那么, 在$\Lambda^p(V)$ 上可以用公式

$$q^{\wedge 0} = 1, \quad q^{\wedge p}(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p) = \begin{vmatrix} f(\mathbf{x}_1, \mathbf{x}_1) & \cdots & f(\mathbf{x}_1, \mathbf{x}_p) \\ & \cdots\cdots\cdots & \\ f(\mathbf{x}_p, \mathbf{x}_1) & \cdots & f(\mathbf{x}_p, \mathbf{x}_p) \end{vmatrix}$$

引出二次型$q^{\wedge p}$.

证明, 在代数$\Lambda(V)$上得到二次型q的延伸也是个非退化的二次型.

5. 对于实的n维欧几里得向量空间$(V, (*|*))$, 称满足条件$(d|d)^{\wedge n} = 1$的元素$d \in \Lambda^n(V)$是这个空间的一个**定向**.

进一步, 设$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$, $(\mathbf{e}'_1, \cdots, \mathbf{e}'_n)$是空间$V$的两个基底, 如果由一个基底向另一个基底的转换矩阵的行列式是正的, 就说这两个基底是**被相同定向的**. 显然, 这个时候基底向量的顺序也是十分紧要的. 同样地显然, 所有有序的基底的集合就被精确地分成两大类, 每一类由被相同定向的有序基底组成, 而来自不同的定向类的基底是**变向的**(或**反向的**). 选择这两类中的一个, 把它称为空间V的**定向**.

在两个定向的定义之间存在某个联系吗?

第7章　附　　录

针对读者范围的广泛性而收集到这一章里的材料, 在线性代数和几何学的基础教程中, 通常是不会完全都给予阐述的, 其明显的原因是缺乏足够的教学时间. 与此同时, 哪怕是了解附录中某些由正文扩展出来的数学资料也是有益的. 作者所在的大学的实践表明, 这类必要的, 最小程度上的超范围的出路会发展成自然的好奇心. 出于同样的效果的考虑, 在本章最后一节介绍了某些没解决的问题. 实际上, 它们还可以更广泛得多: 各式各样的创造性的数学活动中会源源不断地产生待解决的问题.

§1　线性算子的范数与函数

1. 线性算子的范数　设V是域$\mathbb{R}$或域$\mathbb{C}$上的可度量(在第3章§2第5目的意义下)的向量空间, 不一定是有限维的. 如果$||*||$是V上的范数, 那么, 自然地令人感兴趣的是, V上线性算子的向量空间$\mathcal{L}(V)$是否也是可度量的. 至少有两种方式在$\mathcal{L}(V)$上引进范数.

定义1　*把长度为1的向量$\mathbf{v}\in V$上的函数值$\mathbf{v}\to||\mathcal{A}\mathbf{v}||$的上确界*

$$||\mathcal{A}||_s=\sup_{||\mathbf{v}||=1}||\mathcal{A}\mathbf{v}|| \tag{1}$$

*称为线性算子$\mathcal{A}:V\to V$的***模**, *记为$||\mathcal{A}||_s$.*

一般说来, $||\mathcal{A}||_s$的存在性是不能保证的.

例1　在以$||f||=\sqrt{\int_0^1 f(t)^2dt}$为范数的所有多项式作成的空间$V=\mathbb{R}[t]$上, 研究微分算子$\mathcal{D}_t$. 我们有

$$||\sqrt{2n+1}t^n||=\sqrt{(2n+1)\int_0^1 t^{2n}dt}=1.$$

令我们感兴趣的是量

$$||\mathcal{D}_t(\sqrt{2n+1}t^n)|| = ||n\sqrt{2n+1}t^{n-1}|| = n\sqrt{2n+1}\sqrt{\int_0^1 t^{2n-2}dt} = n\sqrt{\frac{2n+1}{2n-1}},$$

它随着多项式的幂次n无限增长.

定理1 如果V是个有限维的向量空间, 那么, 上确界$||\mathcal{A}||_s$必存在.

证明 根据范数的定义和线性算子的定义, 当复合两个连续函数$\mathbf{v} \mapsto \mathcal{A}\mathbf{v}$和$\mathcal{A}\mathbf{v} \to ||\mathcal{A}\mathbf{v}||$时, 函数$\mathbf{v} \mapsto ||\mathcal{A}\mathbf{v}||$必然是向量$\mathbf{v}$的坐标$\alpha_1, \cdots, \alpha_n$的连续函数, $n = \dim V$. 由已知的波尔查诺–魏尔斯特拉斯定理, 连续函数$\mathbf{v} \to ||\mathcal{A}\mathbf{v}||$在有界闭集上(我们这里是$n-1$维的球面$||\mathbf{v}|| = 1$)是个有界函数, 从而上确界$||\mathcal{A}||_s$存在, 因为有界集必有上确界. 进一步, 可以在球面$||\mathbf{v}|| = 1$上找到一点(向量$\mathbf{v}_0$), 在此处$||\mathcal{A}\mathbf{v}||$达到自己的上界. □

这样一来, 就刚好证明了, 至少存在一个单位向量$\mathbf{v}_0 \in V$使得$||\mathcal{A}||_s = ||\mathcal{A}\mathbf{v}_0||$. 具有这种性质的向量有时被称为线性算子$\mathcal{A}$的一个**极大向量**.

已经被赋予范数(1)的有限维空间$\mathcal{L}(V)$在这种论述的通常意义之下就构成了一个可度量空间(显然, 是个巴拿赫空间). 事实上

$$||\mathcal{A}||_s = 0 \Leftrightarrow ||\mathcal{A}\mathbf{v}|| = 0 \Leftrightarrow \mathcal{A}\mathbf{v} = 0\ (\text{对所有满足条件}||\mathbf{v}|| = 1\text{的}\mathbf{v})$$

$$\Leftrightarrow \mathcal{A}\mathbf{v} = 0\ (\text{对所有}\mathbf{v} \in V) \Leftrightarrow \mathcal{A} = \mathcal{O}.$$

进而,

$$||\lambda\mathcal{A}||_s = \sup_{||\mathbf{v}||=1} ||\lambda\mathcal{A}\mathbf{v}|| = \sup_{||\mathbf{v}||=1} |\lambda| \cdot ||\mathcal{A}\mathbf{v}|| = |\lambda| \sup_{||\mathbf{v}||=1} ||\mathcal{A}\mathbf{v}|| = |\lambda| \cdot ||\mathcal{A}||_s,$$

$$||\mathcal{A}+\mathcal{B}||_s = \sup_{||\mathbf{v}||=1} ||(\mathcal{A}+\mathcal{B})\mathbf{v}|| = \sup_{||\mathbf{v}||=1} ||\mathcal{A}\mathbf{v} + \mathcal{B}\mathbf{v}|| \leqslant$$
$$\leqslant \sup_{||\mathbf{v}||=1} (||\mathcal{A}\mathbf{v}|| + ||\mathcal{B}\mathbf{v}||) = \sup ||\mathcal{A}\mathbf{v}|| + \sup ||\mathcal{B}\mathbf{v}|| = ||\mathcal{A}||_s + ||\mathcal{B}||_s.$$

于是, 关于范数定义的所有性质都得到满足, 再进一步, 我们有不等式

$$||\mathcal{A}\mathbf{v}|| \leqslant ||\mathcal{A}||_s \cdot ||\mathbf{v}|| \tag{2}$$

对每个向量$\mathbf{v} \in V$都成立. 实际上,

$$\left\|\frac{\mathbf{v}}{||\mathbf{v}||}\right\| = 1,\ \frac{1}{||\mathbf{v}||}||\mathcal{A}\mathbf{v}|| = \left\|\frac{1}{||\mathbf{v}||}\mathcal{A}\mathbf{v}\right\| = \left\|\mathcal{A}\frac{\mathbf{v}}{||\mathbf{v}||}\right\| \leqslant ||\mathcal{A}||_s.$$

不等式(2)可以当成是范数的另一个定义的基础.

定义2 称以$||*||$为范数的向量空间V上的线性算子$\mathcal{A}$为**有界的**, 如果存在非负实数N使得

$$||\mathcal{A}\mathbf{v}|| \leqslant N \cdot ||\mathbf{v}|| \tag{3}$$

对每个向量$\mathbf{v} \in V$都成立. 使得不等式(3)成立的所有常数的下限$\inf N$就称为是$\mathcal{A}$的**范数**, 并用符号$\|\mathcal{A}\|_i$代表.

上面介绍的例1说明, 算子$\mathcal{D}_t$就不是有界的, 同时, 对有限维情形, 用不等式(2)可以推出不等式(3)且$N \leqslant \|\mathcal{A}\|_s$. 于是, 就有下面的

定理2 在有限维的可度量向量空间V上的每个线性算子$\mathcal{A}$都是有界的.

我们已经看到, $\|\mathcal{A}\|_i \leqslant \|\mathcal{A}\|_s$. 另一方面, 当$||\mathbf{v}|| = 1$时,

$$||\mathcal{A}\mathbf{v}|| \leqslant ||\mathcal{A}||_i \cdot ||\mathbf{v}|| \Rightarrow ||\mathcal{A}\mathbf{v}|| \leqslant ||\mathcal{A}||_i,$$

从而$\|\mathcal{A}\|_s \leqslant \|\mathcal{A}\|_i$. 这样一来, $\|\mathcal{A}\|_i = \|\mathcal{A}\|_s$, 也就是说, 两个范数是同一个. 在下面, 我们把空间V上的范数和线性算子$\mathcal{A} : V \to V$的范数用同一个符号$|| * ||$来代表. 这不会带来误解, 因为对于向量和线性算子, 我们会使用不同的字母.

例2 继续例1, 在空间$\mathbb{R}[t]$上看乘以t的算子$\mathcal{F}_t$. 可以相信, 它是有界的. 实际上,

$$||\mathcal{F}\mathbf{f}|| = \sqrt{\int_0^1 (tf(t))^2 dt} = \sqrt{\int_0^1 t^2 f(t)^2 dt} \leqslant \sqrt{\int_0^1 f(t)^2 dt} = ||f||,$$

所以$\|\mathcal{F}_t\| \leqslant 1$.

注意, 恒等算子$\mathcal{E} : V \to V$的范数等于1, 这是因为, 对所有的$\mathbf{v} \in V$都有$||\mathcal{E}\mathbf{v}|| = ||\mathbf{v}||$. 可以提供一个更显然的例子, 作用在$n$维欧几里得空间上(以$||\mathbf{v}|| = \sqrt{(\mathbf{v}|\mathbf{v})}$为范数)的, 由关系式

$$\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i, \ \ 1 \leqslant i \leqslant n$$

定义的线性算子$\mathcal{A}$是可对角化的, 其中$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$是$V$的一个正交基底. 不失一般性, 我们认为

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n.$$

那么, 对任意一个向量$\mathbf{v} = \sum\limits_i \alpha_i \mathbf{e}_i$, 我们有

$$\begin{aligned} ||\mathcal{A}\mathbf{v}||^2 &= (\mathcal{A}\mathbf{v}|\mathcal{A}\mathbf{v}) \\ &= \sum_{i,j} \alpha_i \alpha_j (\mathcal{A}\mathbf{e}_i|\mathcal{A}\mathbf{e}_j) = \sum_{i,j} \alpha_i \alpha_j (\lambda_i \mathbf{e}_i|\lambda_j \mathbf{e}_j) \\ &= \sum_i \alpha_i^2 \lambda_i^2 \leqslant \lambda_1^2 \sum_i \alpha_i^2 = \lambda_1^2 \cdot ||\mathbf{v}||^2, \end{aligned}$$

从而$||\mathcal{A}\mathbf{v}|| \leqslant |\lambda_1| \cdot ||\mathbf{v}||$, 并且进一步, $||\mathcal{A}|| \leqslant |\lambda_1|$. 因为$||\mathcal{A}\mathbf{e}_1|| = ||\lambda_1 \mathbf{e}_1|| = |\lambda_1|$, 那么, 实际上, $||\mathcal{A}|| = |\lambda_1|$且$\mathbf{e}_1$就是$\mathcal{A}$的一个极大向量.

要补充一个关于算子范数的重要性质, 它与算子复合作用紧密相关.

定理3 设$\mathcal{A}, \mathcal{B}$是可度量空间V上的两个有界算子. 那么, $\mathcal{A}\mathcal{B}$也是个有界算子, 而且

$$||\mathcal{A}\mathcal{B}|| \leqslant ||\mathcal{A}|| \ ||\mathcal{B}||. \tag{4}$$

证明 实际上, 由不等式(2), 对任意向量$\mathbf{v} \in V$, 我们有

$$||(\mathcal{AB})\mathbf{v}|| = ||\mathcal{A}(\mathcal{B}\mathbf{v})|| \leqslant ||\mathcal{A}|| \cdot ||\mathcal{B}\mathbf{v}|| \leqslant ||\mathcal{A}|| \cdot ||\mathcal{B}|| \cdot ||\mathbf{v}||,$$

从而导出算子$\mathcal{AB}$的有界性, 而且依据范数的第二种定义, 它一定满足不等式(4).

推论 设$\mathcal{A} : V \to V$是一个有界算子, 那么

$$||\mathcal{A}^m|| \leqslant ||\mathcal{A}||^m.$$

我们把$\mathcal{L}(V)$上范数的性质再一次列出:

$$\begin{aligned}
&||\mathcal{A}|| = \sup_{||\mathbf{v}||=1} ||\mathcal{A}\mathbf{v}|| = \inf\{N| \ ||\mathcal{A}\mathbf{v}|| \leqslant N \cdot ||\mathbf{v}||\}, \\
&||\lambda\mathcal{A}|| = |\lambda| \cdot ||\mathcal{A}||, \\
&||\mathcal{A}|| = 0 \Leftrightarrow \mathcal{A} = \mathcal{O}, \\
&||\mathcal{A} + \mathcal{B}|| \leqslant ||\mathcal{A}|| + ||\mathcal{B}||, \\
&||\mathcal{A}^m|| \leqslant ||\mathcal{A}||^m.
\end{aligned} \tag{5}$$

不等式$||\mathcal{AB}|| \leqslant ||\mathcal{A}|| \ ||\mathcal{B}||$可以是严格的, 例如, 如果$V = \mathbb{R}^2, \mathcal{P}_x$是平面$\mathbb{R}^2$上点在$x$坐标轴的投影, $\mathcal{P}_y$是在y坐标轴上的投影, 那么

$$\mathcal{P}_x\mathcal{P}_y = \mathcal{P}_y\mathcal{P}_x = \mathcal{O},$$

而此时$||\mathcal{P}_x|| = ||\mathcal{P}_y|| = 1$.

2. 线性算子(矩阵)的函数 前面我们已经看到研究算子$\mathcal{A}$的多项式$f(\mathcal{A})$的必要性. $\mathcal{L}(V)$上范数的存在允许直接把函数级数理论搬到在V上取值的函数(线性算子)上来. 此时, 空间$\mathcal{L}(V)$具有完备性. 如果由线性算子$\mathcal{A}_i$组成的级数$\sum\limits_{i\geqslant 0} \mathcal{A}_i$收敛, 它的和是$\mathcal{A}$, 那么, 在$V$的任意一个基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下, 算子$\mathcal{A}$的矩阵$A$是算子$\mathcal{A}_i$的矩阵$A_i$之和, $A = \sum\limits_{i\geqslant 0} \mathcal{A}_i$. 反过来, 每个收敛的矩阵级数都能对应线性算子的收敛级数.

如同通常的数的函数一样, 我们特别地关注可以表示成幂级数形式的算子函数. 按照范数收敛性的定义, 再由常说的魏尔斯特拉斯判别法, 可以推出, 如果数字级数$\sum\limits_{i\geqslant 0} \alpha_i a^i$收敛, 那么, 在集合

$$\Omega \subset \mathcal{L}(V), \Omega = \{\mathcal{A}| \ ||\mathcal{A}|| \leqslant a\}$$

上的级数$\sum\limits_{i\geqslant 0} \alpha_i \chi^i$必然绝对一致收敛.

例3 如果$\mathcal{A}$是个幂零算子($\mathcal{A}^m = \mathcal{O}$). 那么, 算子$\mathcal{E} - \mathcal{A}$是可逆的, 而且

$$(\mathcal{E} - \mathcal{A})^{-1} = \mathcal{E} + \mathcal{A} + \cdots + \mathcal{A}^{m-1},$$

这很容易直接验证. 现在设V是任意一个巴拿赫空间, 且$\mathcal{A}: V \to V$是一个线性算子, 它的模$||\mathcal{A}|| < 1$. 那么, 同样地, $\mathcal{E} - \mathcal{A}$在$\mathcal{L}(V)$中是可逆的, 且$(\mathcal{E} - \mathcal{A})^{-1} = \sum\limits_{i \geqslant 0} \mathcal{A}^i$. 实际上, 级数$\sum\limits_{i \geqslant 0} \mathcal{A}^i$的收敛性可由几何级数$\sum\limits_{i \geqslant 0} ||\mathcal{A}||^i$的收敛性推出来(把$\mathcal{A}^\circ$理解为恒等算子$\mathcal{E}$), 而它又是模$\sum\limits_{i \geqslant 0} ||\mathcal{A}^i||$的优化级数. 设$\mathcal{B}$是级数$\sum\limits_{i \geqslant 0} \mathcal{A}^i$的和. 那么, $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$就是级数$\sum\limits_{i \geqslant 1} \mathcal{A}^i$的和. 这意味着

$$\mathcal{B}(\mathcal{E} - \mathcal{A}) = (\mathcal{E} - \mathcal{A})\mathcal{B} = \mathcal{E},$$

也就是说, $\mathcal{B}$就是$\mathcal{E} - \mathcal{A}$的逆算子.

3. 指数函数 对于每个线性算子$\mathcal{A}: V \to V$ (V是有限维的可度量空间或者V是个巴拿赫空间, $\mathcal{A}$是有界算子), 可以无条件地说, 最重要的按模收敛的级数是$\sum\limits_{k=0}^{\infty} \frac{1}{k!}\mathcal{A}^k$. 它的收敛性可以得到保障; 范数级数$\sum\limits_{k \geqslant 0} \frac{1}{k!}||\mathcal{A}^k||$可以用收敛级数

$$\sum_{k \geqslant 0} \frac{1}{k!}||\mathcal{A}||^k = \exp ||\mathcal{A}||$$

强化, 后者就是通常的实变量的指数函数.

定义3 级数和$\sum\limits_{k \geqslant 0} \frac{1}{k!}\mathcal{A}^k$ (相应地, $\sum\limits_{k \geqslant 0} \frac{1}{k!}A^k$)用符号$\exp \mathcal{A}$ (相应地, $\exp A$)代表, 被称之为线性算子$\mathcal{A}$的**指数函数**(矩阵A的指数函数). 经常用$e^{\mathcal{A}}$代替$\exp \mathcal{A}$. 由于级数在集合$\Omega = \{\mathcal{A} \mid ||\mathcal{A}|| \leqslant a\}$上的一致收敛性, 函数$\mathcal{X} \mapsto \exp \mathcal{X}$是$\mathcal{L}(V)$到其本身上的一个连续映射.

可以直接在某一行上证明$n \times n$矩阵A的指数函数的存在性. 我们记

$$(a_{ij}) = A, \left(a_{ij}^{(s)}\right) = A^s.$$

设m是对任意参数a_{ij}的绝对值的一个上界: $|a_{ij}| \leqslant m, 1 \leqslant i, j \leqslant n$. 我们可以知道

$$|a_{ij}^{(s)}| \leqslant (nm)^s, \quad 1 \leqslant i, j \leqslant n. \qquad (*)$$

于是, 有

$$|a_{ij}^{(s+1)}| = \left|\sum_{k=1}^{n} a_{ik}^{(s)} a_{kj}\right| \leqslant n(nm)^s m = (nm)^{s+1},$$

也就是说, 不等式(*)对所有的s都成立. 这意味着, 对n^2个对(i, j), 级数$\sum\limits_{s=0}^{\infty} \frac{1}{s!} a_{ij}^{(s)}$都在系数的极大值不超过$m$的矩阵的集合上一致收敛. 级数

$$E + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

是n^2个一致收敛的数字级数并在一起, 本身必然是一致收敛的.

例4 设V是变量t的次数$\leqslant n-1$的实系数多项式的空间，$\mathcal{D}_t$是微分算子. 著名的关于多项式的泰勒公式表明，如果$f(t+a)$是在$f(t)$中用$t+a$代替t得到的次数$\leqslant n-1$的多项式，$a \in \mathbb{R}$. 那么

$$f(t+a) = f(t) + \frac{a}{1!}f^{(1)}(t) + \cdots + \frac{a^{n-1}}{(n-1)!}f^{(n-1)}(t).$$

事实上，可以把它写成

$$f(t+a) = \{\mathcal{E} + \frac{1}{1!}a\mathcal{D}_t + \frac{1}{2!}(a\mathcal{D}_t)^2 + \cdots + \frac{1}{(n-1)!}(a\mathcal{D}_t)^{n-1}\}f(t) = (\exp a\mathcal{D}_t)f(t)$$

的形式.

由于V上算子$a\mathcal{D}_t$的幂零性，级数$\exp a\mathcal{D}_t$可于中间停止而退化成为一个有限和. 这是幂零算子的一个共有的性质. 在我们这个情形，$\exp a\mathcal{D}_t = \mathcal{H}_a$就是用$a \in \mathbb{R}$作用的位置算子，它把$f(t)$变成$f(t+a)$. 它的线性性质可以直接验证.

正如我们已经知道的，在基底

$$\left(\mathbf{e}_i = \frac{t^i}{(n-i)!}\middle| 1 \leqslant i \leqslant n\right)$$

之下，算子$\mathcal{D}_t$对应矩阵$D=J_n$，即n阶的上若尔当块，从而$\mathcal{H}_a$在这个基底之下的矩阵就是

$$H_a = \exp aJ_n = \begin{pmatrix} 1 & a & a^2/2! & \cdots & a^{n-1}/(n-1)! \\ 0 & 1 & a & \cdots & a^{n-2}/(n-2)! \\ 0 & 0 & 1 & \cdots & a^{n-3}/(n-3)! \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

我们应该注意，在指数的定义中，已经认为基础域是实数域或者复数域. 这样，在复变函数中，认为

$$e^{x+iy} = e^x(\cos y + i\sin y) = e^x \cdot e^{iy}, i = \sqrt{-1}.$$

这里包含了欧拉公式

$$e^{iy} = \cos y + i\sin y,$$

它可以利用级数

$$e^{iy} = \sum_{k\geqslant 0}\frac{(iy)^k}{k!} = a(y) + ib(y)$$

以及把函数$\cos y$, $\sin y$分解成关于实变量y的幂级数得到.

例5 复平面$\mathbb{C}$，用域$\mathbb{C}$上的向量空间的观点来看，更好是称它为**复直线**$\mathbb{C}'$，它可以变成实平面$\mathbb{R}^2$. 线性算子$\mathcal{A}: V \to V$, $\mathbf{v} \mapsto (x+y)\mathbf{v}$在基底$(\mathbf{e}_1 = 1, \mathbf{e}_2 = i)$之下对应

矩阵A必形如

$$A=\begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

由复数的指数的定义, 得

$$\exp A=\begin{pmatrix} e^x\cos y & -e^x\sin y \\ e^x\sin y & e^x\cos y \end{pmatrix}=e^x\cdot\begin{pmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{pmatrix}.$$

结论: *把乘以复数$z=x+iy$的乘法算子记为$\mathcal{A}$, 则$\exp\mathcal{A}$刚好是旋转y角再压缩e^x倍.*

利用绝对收敛的级数的乘积中的加项可以被置换这样一个事实, 我们可以得到下面的一个有用的论断

定理4 *如果$n\times n$阶矩阵A和B是可交换的, 那么*

$$(\exp A)(\exp B)=\exp(A+B).$$

证明 我们有

$$\begin{aligned}&(\exp A)(\exp B)\\&=\left(\sum_{s\geqslant0}\frac{1}{s!}A^s\right)\left(\sum_{t\geqslant0}\frac{1}{t!}B^t\right)=\sum_{s,t\geqslant0}\frac{1}{s!}\cdot\frac{1}{t!}A^sB^t=\sum_{k\geqslant0}\sum_{s=0}^{k}\frac{1}{s!(k-s)!}A^sB^{k-s}\\&=\sum_{k\geqslant0}\frac{1}{k!}\left(\sum_{s=0}^{k}\frac{k!}{s!(k-s)!}A^sB^{k-s}\right)=\sum_{k\geqslant0}\frac{1}{k!}\left(\sum_{s=0}^{k}\begin{pmatrix}k\\s\end{pmatrix}A^sB^{k-s}\right)\\&=\sum_{k\geqslant0}\frac{1}{k!}(A+B)^k=\exp(A+B).\end{aligned}$$

在这里, 我们利用了牛顿二项式公式

$$\sum_{s=0}^{k}\begin{pmatrix}k\\s\end{pmatrix}A^sB^{k-s}=(A+B)^k$$

在任意一个交换环, 特别地在环$\mathbb{Z}[A, B]$上成立这一事实.

我们的级数逐项乘积的规律性可用归纳法证明, 事实上.

$$\sum_{k=0}^{2K}\frac{(A+B)^k}{k!}=\left(\sum_{k=0}^{K}\frac{A^k}{k!}\right)\left(\sum_{k=0}^{K}\frac{B^k}{k!}\right)+R_K,$$

其中$R_K=\sum\dfrac{A^s}{s!}\dfrac{B^t}{t!}$是关于所有的满足性质$\max(s,t)>K$且$s+t\leqslant 2K$的数对$(s,t)$求和. 这样的数对共有$K(K+1)$个. 如果$m$是矩阵$A$和$B$的系数的模的一个上界, 那么, 乘积$\dfrac{A^s}{s!}\dfrac{B^t}{t!}$不会超过

$$\frac{n(nm)^s}{s!}\frac{n(nm)^t}{t!}\leqslant\frac{(nm_0)^{2K}}{K!},$$

其中m_0是某个正数. 这就意味着, 矩阵R_K的系数按绝对值不超过

$$\frac{K(K+1)(nm_0)^{2K}}{K!},$$

也就是说, 当K无限增长时, R_K趋于零, 这就给出了我们需要的公式. □

对于零矩阵O, 按定义, 我们有$\exp O=E$. 因为矩阵A和$-A$显然是可换的, 所以, 依据定理4, 有

$$\exp A\cdot\exp(-A)=\exp(A-A)=\exp O=E,$$

由此, 可得到

推论 设A是任意一个$n\times n$阶矩阵(实的或者复的). 那么, $\exp A$是非退化矩阵, 且

$$(\exp A)^{-1}=\exp(-A).$$

4. 线性群的单参数子群 函数exp在很多方面都是饶有兴趣的, 正如可以在下面简单的命题中看到的那样, 有些问题可以用群论的观点去看.

定理5 对任意矩阵$A\in M_n(\mathfrak{K}), \mathfrak{K}=\mathbb{R}$或$\mathfrak{K}=\mathbb{C}$, 由

$$\varphi_A: t\mapsto\exp(tA),\quad t\in\mathbb{R}$$

确定的对应是实数加法群$\mathbb{R}^+$到$GL_n(\mathfrak{K})$的一个连续同态映射.

证明 事实上, 略加改动地引用定理4的推论, 我们容易相信, $\varphi_A(t)\in GL_n(\mathfrak{K})$, 而且

$$(\varphi_A(t))^{-1}=\varphi_A(-t).$$

此外,

$$\varphi_A(s)\cdot\varphi(t)=\exp(sA)\cdot\exp(tA)=\exp(sA+tA)=\exp((s+t)A)=\varphi_A(s+t),$$

也就是说, φ_A是$\mathbb{R}^+$到$GL_n(\mathfrak{K})$的同态, 它的连续性在指数的定义中已经打下了基础. □

定义4 通常称集合$\{\exp(t\mathcal{A})|t\in\mathbb{R}\}$是$V$上**单参数线性算子群**(在一个固定的基底之下就是**单参数矩阵群**).

例6 如果$A=a$是个实数, 那么$\{\exp(ta)|t\in\mathbb{R}\}$就是实数乘法群. 如果$A=\sqrt{-1}=i$, 那么$\{\exp(ti)|t\in\mathbb{R}\}$就是圆周.

为了得到关于矩阵$\exp A$(算子$\exp\mathcal{A}$)的本性的更加确切的结论, 我们证明下面的命题.

定理6 设$\lambda_1,\lambda_2,\cdots,\lambda_n$是矩阵$A\in M_n(\mathbb{C})$的特征根, 这里是连同重数列出的(也就是说, λ_i不一定是不同的).

那么, 矩阵$\exp A$的特征根恰好是

$$\exp\lambda_1,\exp\lambda_2,\cdots,\exp\lambda_n.$$

证明 正如我们已经知道的(第2章§4的定理1), 每个复矩阵都相似于一个三角形矩阵. 特别地, 对我们这里的矩阵A可找到一个非退化矩阵B使得$A = B^{-1}TB$, 其中T是个(上)三角形矩阵. 我们同样知道, A和T有相同的特征根. 因为三角形矩阵的特征根就是它的主对角线上的元素, 所以, 不失一般性, 可以认为, 它们按顺序就是$\lambda_1, \cdots, \lambda_n$. 这样一来,

$$T = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

我们有

$$T^k = \begin{pmatrix} \lambda_1^k & * & \cdots & * \\ 0 & \lambda_2^k & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix},$$

也就是

$$\exp T = \begin{pmatrix} \exp\lambda_1 & * & \cdots & * \\ 0 & \exp\lambda_2 & \cdots & * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \exp\lambda_n \end{pmatrix},$$

这说明$\exp\lambda_1, \cdots, \exp\lambda_n$是矩阵$\exp T$的特征根. 只要注意到$\exp A = B^{-1}\cdot\exp T\cdot B$ (见习题7.1.1), 就进一步知道$\exp\lambda_i$同样是矩阵A的特征根. □

推论 对任意方阵A, 关系式

$$\det\exp A = \exp(\operatorname{tr} A) \tag{6}$$

成立.

证明 事实上, 矩阵的行列式等于它的所有的特征根的乘积, 而迹等于它的所有的特征根之和, 因为这些都可以从定理6直接得到. 更显然地,

$$\begin{aligned}\det\exp A &= \det(B^{-1}\cdot\exp T\cdot B) = \det\exp T \\ &= \prod_{i=1}^{n}\exp\lambda_i = \exp(\lambda_1 + \cdots + \lambda_n) = \exp(\operatorname{tr}A).\end{aligned}$$ □

由关系式(6)可以再次推出, $\exp A$总是一个非退化矩阵, 而且, 对所有的$A \in M_n(\mathbb{R})$, $\det\exp A$总是一个正实数.

对于矩阵(线性算子)的指数, 我们还可做若干说明. 一如既往, 用矩阵${}^tA$代表矩阵A的转置矩阵, $\overline{A}$代表A的元素用自己的共轭数替换之后得到的矩阵(如果A是个实矩阵, 那么, $\overline{A} = A$). 我们再回顾

$${}^t(A^k) = ({}^tA)^k, \quad \overline{A^k} = (\overline{A})^k. \tag{7}$$

关于第一个关系式可看[BA I]第2章§3的第3目, 而第二个关系式可以由这样的一个简单的事实推出来, 即矩阵A^k的系数可以记成矩阵A的系数的k次齐次型, 映射$z \mapsto \overline{z}$是域$\mathbb{C}$上的自同构. 进而,

$$ {}^t(A+B) = {}^tA + {}^tB, \quad \overline{A+B} = \overline{A} + \overline{B}. $$

于是, 由$\exp A$的定义和关系式(7), 我们直接可以得到

$$ {}^t(\exp A) = \exp({}^tA), \quad \overline{\exp A} = \exp \overline{A}. \tag{8} $$

下面, 我们使用对某些矩阵空间的公认的记法:

i) $\mathfrak{gl}(n,\mathfrak{K}) = \mathfrak{gl}_n(\mathfrak{K}) = M_n(\mathfrak{K})$是域$\mathfrak{K} = \mathbb{C}$或者$\mathfrak{K} = \mathbb{R}$上的所有$n$阶矩阵的空间;

ii) $\mathfrak{sl}(n,\mathfrak{K}) = \mathfrak{sl}_n(\mathfrak{K})$是所有迹为零的$n \times n$阶矩阵的空间;

iii) $\mathfrak{so}(n,\mathfrak{K}) = \mathfrak{so}_n(\mathfrak{K})$是所有反交换的$n \times n$阶矩阵(实的或者复的), ${}^tX + X = 0$;

iv) $\mathfrak{u}(n) = \mathfrak{u}_n$是$\mathbb{R}$上(但不是$\mathbb{C}$上)的斜埃尔米特矩阵$X \in M_n(\mathbb{C})$, $X^* + X = 0$ $(X^* := {}^t\overline{X})$作成的空间;

v) $\mathfrak{su}(n) = \mathfrak{su}_n$ 是$\mathbb{R}$上所有迹为零的斜埃尔米特对称n阶矩阵作成的空间.

在第2章§2的例6中, 我们已经引进了李代数的概念(同样地可见第3章§3的例2). 下列命题是真确的.

命题　空间i)—v)的每一个, 对于矩阵的换位算子$[A,B] = AB - BA$都作成一个(古典的)李代数.

证明　基于在第1章我们已经熟知的映射

$$ X \mapsto {}^tX, \quad X \mapsto X^*, \quad X \mapsto \mathrm{tr}X $$

的基本性质就可以做初等的验证. 以iv)为例,

$$ \begin{aligned} [A,B]^* &= (AB - BA)^* \\ &= (AB)^* - (BA)^* = B^*A^* - A^*B^* = (-B)(-A) - (-A)(-B) \\ &= BA - AB = -[A,B]. \end{aligned} $$

在情形v), 添加在斜对称之下保持$\mathrm{tr}X = 0$不变的条件, 从而必恒有

$$ \mathrm{tr}[X,Y] = \mathrm{tr}AB - \mathrm{tr}BA = 0. \qquad \square $$

再次约定一些符号, 这次我们屡屡见到的都是群(典型群):

i′) $GL(n,\mathfrak{K}) = GL_n(\mathfrak{K})$是$\mathfrak{K} = \mathbb{R}$或者$\mathfrak{K} = \mathbb{C}$上的$n$阶完全线性群;

ii′) $SL(n,\mathfrak{K}) = SL_n(\mathfrak{K})$是$M_n(\mathfrak{K})$中行列式为1的矩阵作成的特殊线性群;

iii′) $SO(n) = SO(n,\mathbb{R}) = SO_n(\mathbb{R})$ (或者$SO(n,\mathbb{C})$)是所有实的或者复的行列式为1的正交矩阵作成的群;

iv′) $U(n) = U_n$是n阶酉矩阵作成的群;

v′) $SU(n)=SU_n$是(行列式为1的酉矩阵)特殊酉矩阵群.

定理7 映射$X\mapsto \exp X$建立了已指明的类型的i)型李代数和i′)型群之间的关系式, $1\leqslant i\leqslant 5$.

证明 i)可以看定理4的推论.

ii) 断言在i)和公式(6)中已被确认.

iii) 如果${}^tX+X=0$, 那么, 应用定理4的推论和公式(6), 我们有

$${}^t(\exp X)=\exp{}^tX=\exp(-X)=(\exp X)^{-1},$$

也就是

$${}^t(\exp X)\cdot\exp X=E\Rightarrow \exp X\in SO(n).$$

iv) $X^*+X=0\Rightarrow(\exp X)^*=\exp X^*=\exp(-X)=(\exp X)^{-1}$
$\Rightarrow \exp X\in U(n)$.

v) 如果$X\in\mathfrak{su}(n)$, 那么$\det\exp X=1$, 从而$\exp\ X\in SU(n)$. □

本来还可以试图去证明, 有相互单值确定的同一个类型中的关系式

$$A\in\cdots\Leftrightarrow \exp A\in\cdots,$$

只要在左侧的矩阵的李代数中对任意小的$\varepsilon>0$, 有$\|A\|<\varepsilon$在右侧就可以取到型如$\exp A=E+B$的矩阵, 使$\|B\|<\delta$. 但定理7中断言的箭头不能完全反过来, 正如下例所示

例7 若尔当块

$$J_2(-1)=\begin{pmatrix}-1&1\\0&-1\end{pmatrix}$$

是群$SL(2,\mathbb{C})$的一个元素. 设$J_2(-1)=\exp A$, 其中$A\in\mathfrak{sl}(2,\mathbb{C})$. 因为$J_2(-1)$不能化成对角形, 所以矩阵$A$同样也不能对角化(见习题7. 1. 1). 这样一来, 矩阵A的特征根λ_1和λ_2应当重合, 也就是

$$\lambda_1=\lambda_2,\lambda_1+\lambda_2=0\Rightarrow\lambda_i=0,\ \ i=1,2.$$

但是, $\mu_i:=\exp\lambda_i=\exp 0=1$应该是矩阵$J_2(-1)$的特征根(定理6), 这个时候, 实际上$\mu_i=-1$. 这就意味着矩阵$A$是不存在的. 换言之, 并不是群$SL(2,\mathbb{C})$的每一个元素必然属于一个单参数子群.

5. 谱半径 如果试着去解下面的习题9就会得到这样的结论, 在有限维埃尔米特空间上正规算子$\mathcal{A}$的范数$\|\mathcal{A}\|$等于它的特征值的最大值.

定义5 设$\mathcal{A}$是$\mathbb{C}$上n维向量空间V的任意一个线性算子, $\lambda_1,\cdots,\lambda_n$是它的特征值. 那么, 称

$$r(\mathcal{A})=\max_{j=1,\cdots,n}|\lambda_j|$$

是算子$\mathcal{A}$的谱半径.

正如我们已经看到的, 对于可对角化的线性算子$\mathcal{A}$, $r(\mathcal{A}) = ||\mathcal{A}||$. 但是, 一般说来, $r(\mathcal{A}) \leqslant ||\mathcal{A}||$. 实际上, 如果$\lambda = r(\mathcal{A})$且$\mathbf{v}$是属于$\lambda$的特征向量($||\mathbf{v}|| = 1$), 那么,

$$||\mathcal{A}|| = \sup_{||\mathbf{x}||=1} ||\mathcal{A}\mathbf{x}|| \geqslant ||\mathcal{A}\mathbf{v}|| = ||\lambda\mathbf{v}|| = |\lambda|.$$

这个不等式可以是严格的.

例8　设$(\mathbf{e}_i)$是带有标准纯量乘积(内积)

$$(\mathbf{x}|\mathbf{y}) = \sum_i x_i\overline{y}_i \quad (\mathbf{x} = \sum_i x_i\mathbf{e}_i, \quad \mathbf{y} = \sum_i y_i\mathbf{e}_i)$$

的线性空间V的一个标准正交基底. 线性算子$\mathcal{A}$在这个基底之下对应矩阵$A = J_n(\lambda)$. 为简便计, 我们认为$\mathfrak{K} = \mathbb{R}$, $\lambda \in \mathbb{R}$. 所以,

$$||\mathcal{A}\mathbf{x}|| = \sqrt{(x_1\lambda + x_2)^2 + \cdots + (x_{n-1}\lambda + x_n)^2 + (x_n\lambda)^2}.$$

如果$\| \mathbf{x} \| = \sqrt{x_1^2 + \cdots + x_n^2} = 1$, 那么

$$||\mathcal{A}\mathbf{x}|| = \sqrt{\lambda^2 + 2\lambda\sum_{i=1}^{n-1} x_i x_{i+1} + (1 - x_1^2)},$$

而且经过简单地整理即表明

$$r(\mathcal{A}) = \lambda < \sup_{||\mathbf{x}||=1} ||\mathcal{A}\mathbf{x}|| = ||\mathcal{A}||.$$

其次, 当$K \geqslant n$时, $\lambda = 0 \Rightarrow (J_n(0))^k = 0$, 从而, $r(\mathcal{A}) = 0 = \lim\limits_{k\to\infty} ||\mathcal{A}||^{1/k}$. 把习题2. 4. 12弄明白, 我们会发现, 当$\lambda \neq 0$时(且$||\mathbf{x}|| = 1$)必有等式

$$\begin{aligned}
&||(J_n(\lambda))^k\mathbf{x}||^{1/k}\\
&= \lambda\{(x_1 + \frac{k}{\lambda}x_2 + \cdots + \frac{\binom{k}{n-1}}{\lambda^{n-1}}x_n)^2 + (x_2 + \frac{k}{\lambda}x_3 + \cdots + \frac{\binom{k}{n-2}}{\lambda^{n-2}}x_n)^2 + \cdots\}^{1/2k}\\
&= \lambda(1 + a_1(\mathbf{x},\lambda)k + \cdots + a_{n-1}(\mathbf{x},\lambda)k^{n-1})^{1/(2k)},
\end{aligned}$$

也就是

$$\begin{aligned}
&\lim_{k\to\infty} ||(J_n(\lambda))^k\mathbf{x}||^{1/k}\\
&= \lambda \cdot \lim_{k\to\infty}(1 + a_1(\mathbf{x},\lambda)k + \cdots + a_{n-1}(\mathbf{x},\lambda)k^{n-1})^{1/2k} = \lambda \cdot 1 = r(\mathcal{A}).
\end{aligned}$$

利用我们已经研究过的若尔当标准型理论和例子, 可以指出, 对一般线性算子$\mathcal{A}$的情形, 有下列事实:

i) $r(\mathcal{A}) = \lim\limits_{k\to\infty} ||\mathcal{A}^k||^{1/k}$;

ii) $r(\mathcal{A}) \leqslant ||\mathcal{A}^k||^{1/k} \leqslant ||\mathcal{A}||$ (见定理3的推论).

依据[2]第1章§10的关于有限维空间上模的等价性的定理7(此时必然是个巴拿赫空间), 我们还可以导出不等式

iii) $r(\mathcal{A}) \leqslant \max\limits_k \sum\limits_{j=1}^n |a_{jk}|, r(\mathcal{A}) \leqslant \max\limits_j \sum\limits_{k=1}^n |a_{jk}|$, 对任意矩阵$A \in M_n(\mathbb{C})$都成立.

说明 显然, 在任意基底之下线性算子$\mathcal{A}$对应矩阵A, 必有$r(\mathcal{A}) = r(A)$, 而且所有的与谱半径有关的结果都可以用矩阵语言加以翻译. 关于算子范数的更详尽的性质可以跟踪[15].

习 题

1. 验证

$$\dim \mathfrak{sl}(n, \mathfrak{K}) = n^2 - 1, \dim \mathfrak{so}(n, \mathfrak{K}) = n(n-1)/2, \dim \mathfrak{su}(n) = n^2 - 1.$$

2. 建立群$U(1)$与群$SO(2, \mathbb{R})$之间的同构.

3. 下列矩阵称之为$\mathbb{C}$上的泡利(W. Pauli)矩阵

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

我们令

$$T_1 = \frac{1}{2}i\sigma_1, \quad T_2 = -\frac{1}{2}i\sigma_2, \quad T_3 = \frac{1}{2}i\sigma_3.$$

显然,

$$\langle T_1, T_2, T_3\rangle_{\mathbb{C}} = \mathfrak{sl}(2),$$

$$\langle T_1, T_2, T_3\rangle_{\mathbb{R}} = \mathfrak{su}(2).$$

试验证

$$[T_1, T_2] = T_3, \quad [T_3, T_1] = T_2, \quad [T_2, T_3] = T_1.$$

4. 建立李代数$\mathfrak{so}(3, \mathbb{R})$与3维欧几里得空间关于向量乘积的向量李代数$\mathfrak{su}(2)$之间的同构.

5. 证明, 如果A, B是$n \times n$矩阵且$\det B \neq 0$, 那么,

$$\exp(B^{-1}AB) = B^{-1}(\exp A)B.$$

6. 对于矩阵

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

求出$\exp A$.

7. 证明, 每个酉算子$\mathcal{A}$都可以表示成$\mathcal{A} = e^{i\mathcal{B}}$的形式, 其中$\mathcal{B}$是一个埃尔米特算子.

8. 证明, 如果$\mathcal{A}$是个酉算子, 而$\mathcal{B}$是任意一个线性算子, 那么$||\mathcal{AB}|| = ||\mathcal{B}||$.

9. 对每个规范算子$\mathcal{A}$, 都有$||\mathcal{A}^k|| = ||\mathcal{A}||^k$, $k = 1, 2, 3, \cdots$.

10. 证明, $\lim\limits_{k\to\infty} \mathcal{A}^k = \mathcal{O} \Leftrightarrow r(\mathcal{A}) < 1$.

§2 线性微分方程

1. 指数函数的导数 设$P(t) = (p_{ij}(t))$是个矩阵, 它的系数$p_{ij}(t)$是实变量t的可微函数, 按定义, 我们令

$$\frac{d}{dt}P(t) := \left(\frac{dp_{ij}(t)}{dt}\right),$$

并且称$P'(t) = dP/dt$是矩阵P的导数, 然后就显然地, 可以得到矩阵乘积的微分法则

$$\frac{d}{dt}(PQ) = \frac{dP}{dt}Q + P\frac{dQ}{dt}.$$

一般说来, 正如可以以矩阵$\begin{pmatrix} 0 & t \\ 1 & 0 \end{pmatrix}$为例表明的那样,

$$P'(t)P(t) \neq P(t)P'(t),$$

例如,

$$\frac{dP^2}{dt} \neq 2P\frac{dP}{dt}.$$

但是, 我们对于一个矩阵可与其导数矩阵交换的特殊情形有兴趣.

定理1 设A是个常系数矩阵(实的或者复的), $F(t) = \exp(tA)$.

那么,

$$\frac{d}{dt}F(t) = A \cdot F(t). \tag{1}$$

证明 按其定义, 矩阵A和$F(t)$是可交换的, 也就是说, 要证明的关系式的右端可以$F(t)A$代替. 用Δt代表变量t的微小增量. 按§1的定理4, 我们有

$$F(t+\Delta t) = F(t)F(\Delta t),$$

所以,

$$\begin{aligned}
&\frac{1}{\Delta t}[F(t+\Delta t) - F(t)] \\
&= \frac{1}{\Delta t}[F(\Delta t) - E] \cdot F(t) = \frac{1}{\Delta t}\left[\sum_{i=0}^{\infty}\frac{1}{i!}(\Delta tA)^i - E\right] \cdot F(t) \\
&= \left[\sum_{i=1}^{\infty}\frac{1}{i!}(\Delta t)^{i-1}A^i\right] \cdot F(t).
\end{aligned}$$

由幂级数的连续性, 我们可导出结论

$$\lim_{\Delta t\to 0}\left[A + \frac{1}{2!}(\Delta t)A^2 + \frac{1}{3!}(\Delta t)^2A^3 + \cdots\right] = A.$$

所以,

$$\frac{d}{dt}F(t) := \lim_{\Delta t\to 0}\frac{1}{\Delta t}[F(t+\Delta t) - F(t)] = A \cdot F(t). \qquad \square$$

2. 微分方程 现在设

$$\begin{aligned} &z_1' = a_{11}z_1 + a_{12}z_2 + \cdots + a_{1n}z_n, \\ &z_2' = a_{21}z_1 + a_{22}z_2 + \cdots + a_{2n}z_n, \\ &\cdots\cdots\cdots\cdots \\ &z_n' = a_{n1}z_1 + a_{n2}z_2 + \cdots + a_{nn}z_n \end{aligned} \tag{2}$$

是1阶的常系数的线性微分方程组, 并记为

$$\frac{d\mathbf{z}}{dt} = A\mathbf{z}, \tag{2'}$$

这里$\mathbf{z} = [z_1, z_2, \cdots, z_n]$是向量列;它的分量$z_i = z_i(t), 1 \leqslant i \leqslant n$, 是$t$的未知的可微函数. t在一个区间$r_1 < t < r_2$中, 且满足一个初始条件$z_i(0) = z_i^0$.

微分方程的一般理论[1]要确保$\mathbf{z} = \mathbf{z}(t)$的存在性和唯一性. 此外, 由(2)和(2′)可以看出, 对应不同的初始条件的解构成一个向量空间, 而常说的基本解就是这个空间的一个基底.

重新回到关系式(1), 我们发现, 允许做这样的解释: 矩阵$F(t) = (f_{ij}(t))$的每个列$\mathbf{z}_j = [f_{1j}, \cdots, f_{nj}]$都是方程组(2)的一个解, 它满足初始条件$\mathbf{z}_j(0) = [0, \cdots, 1, \cdots, 0]$, 这里边只有第$j$个位置上是1(因为$F(0) = E$). 因为初始条件在$j = 1, \cdots, n$时是线性无关的, 而其余的都是它们的线性组合, 所以, 矩阵$F(t)$的n个列就取尽了方程组(2)的基本解的整个集合. 解空间是n维的.

更一般的线性微分方程组形如

$$\frac{d\mathbf{z}}{dt} = A\mathbf{z} + \mathbf{b}, \tag{3}$$

其中

$$A = (a_{ij}), \quad \mathbf{z} = [z_1, \cdots, z_n], \quad \mathbf{b} = [b_1, \cdots, b_n].$$

如同在代数方程组的情形一样, 方程组(3)可能是不相容的. 如果不是这样, 且$\mathbf{z}_0$是某一个特解, 那么, 方程组(3)的一般解就是$\widetilde{\mathbf{z}} = \mathbf{z} + \mathbf{z}_0$, 其中$\mathbf{z}$是与(3)相对应的齐次方程组(2)的一般解.

3. n阶线性微分方程 有形如

$$z^{(n)} + a_1 z^{(n-1)} + \cdots + a_{n-1}z^{(1)} + a_n z = 0 \tag{4}$$

的方程, 其中, 依惯例, $z = z(t)$是自变量t的未知函数, 而系数$a_1, \cdots, a_n$是实数或者复数. 先不涉及满足初始条件

$$z(0) = z^0, z^{(1)}(0) = z_1^0, \cdots, z^{(n-1)}(0) = z_{n-1}^0,$$

1) 可以参见, 例如, 庞特里亚金. П.С. 的常微分方程, 中译本, 林武忠、倪明康译, 北京: 高等教育出版社, 2006.

的解$z(t)$的存在性和唯一性理论, 我们注意到, 关于z和它的各级导数的线性方程(4)的解空间的线性性质. 解本身可以通过把(4)化成特殊的方程组(2)而找到, 也就是, 设

$$z_1 = z, \quad z_2 = z^{(1)}, \quad z_3 = z^{(2)}, \cdots, \quad z_{n-1} = z^{(n-2)}, \quad z_n = z^{(n-1)},$$

我们就可以得到一个具有特殊形式的矩阵

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-3} & \cdots & -a_1 \end{pmatrix}$$

的线性方式组(2′). 进而我们就可以使用在第2目中的同样手法了.

现在, 可以从另一个方向走近方程式(4). 我们转向作用在无穷次可微函数的空间上线性算子. 设$\mathcal{D}_t$是对t的微分算子, 且

$$\chi(\mathcal{D}_t) = \mathcal{D}_t^n + a_1\mathcal{D}_t^{n-1} + \cdots + a_{n-1}\mathcal{D}_t + a_n\mathcal{E}.$$

这样, 就可以把(4)改写成

$$\chi(\mathcal{D}_t)z = 0. \tag{4′}$$

称多项式$\chi(t)$是微分方程(4′)的**特征多项式**. 这个术语是有确定的含义的. 试寻求形如$z = e^{\lambda t}$的解, 我们得到关系式

$$\chi(\lambda)e^{\lambda t} = \chi(\mathcal{D}_t)e^{\lambda t} = 0,$$

那么, 可得出

$$e^{\lambda t}\text{是方程(4′)的解} \Leftrightarrow \chi(\lambda) = 0.$$

定理2 设$\lambda_1, \cdots, \lambda_m$是方程(4′)的特征多项式$\chi(t)$的所有的不同的特征根, 同时, λ_j的重数是k_j, 使得$k_1 + \cdots + k_m = n$. 那么, 函数

$$t^k e^{\lambda_j t}, \quad 1 \leqslant j \leqslant m, \ 0 \leqslant k \leqslant k_j - 1,$$

作成一个基础解系.

我们不打算推演这个证明. 介绍读者参看. П.С. 庞特里亚金的教学参考书. 在单根的情形(所有的$k_j = 1$), 推论本身并不复杂, 但有重根时就需要用到若尔当标准型了(见第2章).

我们的任务在于举例说明在最简单类型的微分方程理论中的线性代数学的方法.

§3 凸多面体与线性规划

1. 问题的提出 线性规划的基本问题可以用下面的方式提出来. 在实数域$\mathbb{R}$上有限维仿射空间$\mathbb{A}$上有$m+1$个仿射线性函数$f_1, \cdots, f_m; f: \mathbb{A} \to \mathbb{R}$. 需要找出一个点$\dot{a} \in \mathbb{A}$(或一些点), 满足条件

$$f_1(\dot{a}) \geqslant 0, \ \cdots, \ f_m(\dot{a}) \geqslant 0,$$

且使函数f在这些限制之下取得最大的可能值.

与此相反, 某些不等式取反向, $f_i(\dot{a}) \leqslant 0$, 同时(或者)要求去寻求f取最小的可能值, 这可以归结为对从前已做过的替换相应函数符号的情形. 条件$f_i(\dot{a}) = 0$等价于$f_i(\dot{a}) \geqslant 0$和$-f_i(\dot{a}) \geqslant 0$的总和. 所有的函数都可以认为不恒为常数.

2. 论据 考虑下面的一个生产过程的数学模型. 设有一个企业, 使用m种不同的资源, 生产出n种不同的产品, 资源和产品分别按自己的单位用非负的实数计量(我们不把这些量, 如汽车的数量, 当做整数;当大规模的生产和大量需求资源的时候, 它就接近于"连续"模型了).

给定的企业的所有类型产品的产量记成一个**产品向量**$(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$. 开拓广些, 可以得到下列的资源需求的线性模型.

设标号$i(i=1,2,\cdots,n)$的资源在生产一个单位的$j(j=1,2,\cdots,m)$型产品时的支出是$a_{ij} \geqslant 0$, 同时, 资源i的限制量是b_i. 也就是说

$$f_i(x_1, \cdots, x_n) = b_i - \sum_{j=1}^{n} a_{ij} x_j \geqslant 0, \quad 1 \leqslant i \leqslant m,$$

或者, 等价于

$$\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \quad i = 1, 2, \cdots, m. \tag{1}$$

同时还自然地要求

$$x_j \geqslant 0, \quad j = 1, 2, \cdots, n, \tag{2}$$

也就是说, 企业不谋求指派给它的产品以外的销售品或者储备品. 再假设, 不等式组(1)和(2)是相容的. 任意满足这些不等式条件的产品向量都被称为是**可允许的**.

其次, 设π_j是企业从一个单位的j种产品中获得的利润. 我们设

$$f(x_1, \cdots, x_n) = \sum_{j=1}^{n} \pi_j x_j. \tag{3}$$

在线性规划中就简洁地称函数(3)是**目标函数**. 提供函数值的极大值的可允许产品向量就是按利润**最优化**的产品向量. 企业关心的结论是获得最大的利润, 也就是正常地安排进来的资源获得最优化的产品向量. 我们将会看到, 这个问题是就要在例子中描述的问题的一个特殊情形.

在研究线性规划一般问题的内容的更为丰富的几何解释之前，我们先看一个具体的经济生活中的例子.

例　设某个联合公司(我们称它为MBRB)生产两种产品：摩托艇和客运小轮船(分别称为MB和RB). 公司掌握四种类型的设备，每种都有固定的生产能力：组装MB，组装RB，组装发动机和冲压金属板. 问题是公司应该生产多少摩托艇和多少客运小轮船?从MB或者RB上获得的利润取决于摩托艇和客运小轮船的市场价格以及公司的固定的花费. 我们假设，市场价格和生产的可变成本的平均值(经济学中周知的一个概念)都是不变的. 也就是说，在一个合理的期限内是不变的，更加确定地，我们允许摩托艇的市场价格是70 000卢布，而客运小轮船的价格是125 000卢布. 生产MB的可变成本的平均值是65 500卢布，而生产RB的可变成本的平均值是120 000卢布.

这样一来，公司扣除每件MB产品的可变成本可以获得4 500卢布，而扣除每件RB产品的可变成本可以获得5 000卢布，如果N_{mb} (相应地，N_{rb})是生产公司每天生产MB (相应地RB) 的个数，那么，公司的收益(扣除了固定成本)应该等于

$$\pi = f(N_{mb}, N_{rb}) = 4\,500N_{mb} + 5\,000N_{rb} \tag{4}$$

我们假定，一天内生产中每产出一个MB(相应地RB)要使用15%的MB设备能力，12%的发动机设备能力，9%的金属冲压设备能力(相应地，17%的RB设备能力，8%的发动机设备能力以及13%的金属冲压设备能力). 显然，公司的领导受到如下限制

$$\begin{aligned} 0 \leqslant 15N_{mb} &\leqslant 100, \\ 0 \leqslant 17N_{rb} &\leqslant 100, \\ 12N_{mb} + 8N_{rb} &\leqslant 100, \\ 9N_{mb} + 13N_{rb} &\leqslant 100. \end{aligned}$$

借助图23可以给出明显的解释.

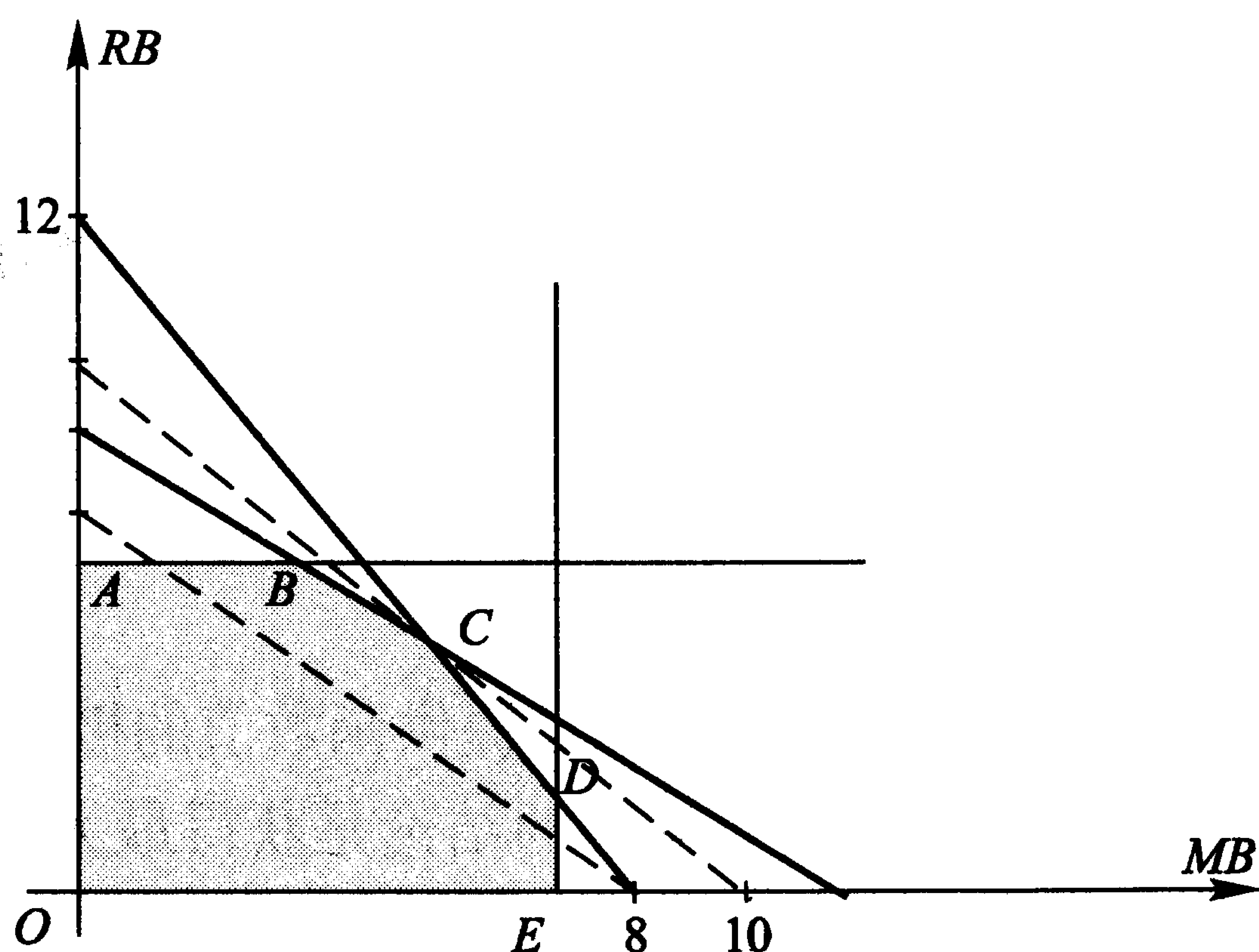


图23

为了清点所有上述强调过的限制, 摩托艇和客运小轮船的产量应该位于区域 $OABCDE$ 内, 这也就是公司领导可以运作的"平台". 同样, 收益的直线(图上的虚线)每一条代表MB 和RB 生产的不同组合, 它们给出的总收益是相同的. 图形显示. 最优解位于C 点, 在这里, 水面运输公司每日生产5. 95MB 和3. 57RB. 按着这种给定的产量生产, MBRB 公司的总收益是每天核计44 625 卢布.

在这个刚讨论过的例子中, 目标函数(4)在凸多边型的顶点C取得极大值. 我们可以郑重地说明, 这绝对不是偶然的. 当然, 线性规划的实际应用与寻求最优生产向量的具体的深入研究的算法有关系. 这可以用人工方法, 也可以在计算机上实现. 这里, 我们只限于应用问题的几何方面, 当然, 它是所有算法的基础.

3. 基本的几何概念 我们在域$\mathbb{R}$上取定一个有限维的仿射空间$\mathbb{A}$. 字母f以及带有指标的f均为$\mathbb{A}$上的仿射线性函数.

称形如

$$\{\dot{a} \in \mathbb{A} | f(\dot{a}) \geqslant 0\}$$

的点的集合是个**半空间**, 其中f是个非常数仿射线性函数. 有限多个半空间交称为**多面体**.

回顾一下(见第4章§3的第6目), 子集$S \subset \mathbb{A}$是凸的, 如果可由$\dot{a}_1, \dot{a}_2 \in S$及$0 \leqslant \lambda \leqslant 1$推出$\lambda\dot{a}_1 + (1-\lambda)\dot{a}_2 \in S$. 由于有

$$f(\lambda\dot{a}_1 + (1-\lambda)\dot{a}_2) = \lambda f(\dot{a}_1) + (1-\lambda) f(\dot{a}_2),$$

所以, 每个半空间都一定是凸的. 又因为任意多个凸集的交集仍然是凸集, 所以, 每个多边形都是凸的. 与从前一样, 我们说任意点

$$\lambda\dot{a}_1 + (1-\lambda)\dot{a}_2, \quad 0 < \lambda < 1,$$

都是以$\dot{a}_1$和$\dot{a}_2$为端点的线段$\dot{a}_1\dot{a}_2$的一个**内点**.

设S是个凸集. 凸的子集$T \subset S$被称为是集合S的一个**界面**, 如果端点在S的线段只要有某些内点含在T中, 那么该线段就整个地含在T中. 整个集合S就是自己的一个界面. 集合S的只由一个点组成的界面被称为是S的一个**顶点**. (读者应该自己能在3维空间中表达出立方体、八面体和多边形的角度, 使得能掌握满足基本要求的直观图形, 这些图形对线性规划理论十分重要. 我们实际上定义的这些图形的界面就是中学平面几何学中边、顶点以及各图形本身. 实心球体的顶点就是它的表面上的所有的点. 多边形S的所有界面的个数不能超过仿射线性函数组$f_1, \cdots, f_m$的子空间的个数. 从而是有限的.)

很自然地, 称多面体$S \subset \mathbb{A}$是**有界的**, 如果对$\mathbb{A}$中任意一个坐标系都能找到一个数N使得任意一点$\dot{a} \in S$的坐标, 按绝对值, 都不超过N. 这个定义与坐标系的选择无关.

下面, 我们要证明一个最重要的结论, 在有界的多面体上的仿射线性函数的极大值(在应用中这种情形最为常见)必将在它的一个顶点上达到;必然是个有限数. 但是, 着手证明之前, 先研究一下多面体及其界面的结果将更为方便.

引理1 凸集S的一族界面的交以及界面的界面都是S的界面.

证明 i) 设$T=\bigcap_i T_i$, T_i都是S的界面. 任意端点在S中的, 内点都属于T_i的线段都必然整个地属于T_i, 这意味着, 它的端点必在T中, 所以它整个地在T中.

ii) 设$T_1\subset T\subset S$, T是集合S的一个界面. 任意端点在S中的线段, 只要内点都在T_1中, 那么必然整个地在T中, 因为T是S的一个界面. 这意味着, 它的端点也在T中, 从而整个地在T_1中, 因为T_1是T的一个界面。 □

引理2 设S是个多面体, 由

$$f_i\geqslant 0,\quad i=1,\cdots,m$$

给定. 那么, 对任意一个指标i, 多面体

$$S_i=S\cap\{\dot a|f_i(\dot a)=0\}$$

或者是空集, 或者是S的一个界面.

证明 设S_i是非空的, $\dot a_1,\dot a_2\in S$且线段$\lambda\dot a_1+(1-\lambda)\dot a_2$的内点在$S_i$中, 函数

$$f_i(\lambda\dot a_1+(1-\lambda)\dot a_2),\quad 0\leqslant\lambda\leqslant 1$$

对λ是线性的, 对某个$0<\lambda_0<1$, 它变成零, 此外, 当$\lambda=0$和$\lambda=1$时, 是非负的, 因此, 它应当恒等于零, 所以, 整个线段都在S_i中. □

引理3 多面体

$$S=\{\dot a|f_i(\dot a)\geqslant 0;\quad 1\leqslant i\leqslant m\}$$

上的非常数的仿射线性函数f, 在那些使得所有$f_i(\dot a)>0$的点$\dot a\in S$上, 不可能取得最大值.

证明 因为f不是常数, 所以$Df\neq 0$. 在与$\mathbb{A}$相伴的向量空间V中选取一个向量$\mathbf{v}\in V$, 使得$Df(\mathbf{v})\neq 0$. 可以认为$Df(\mathbf{v})>0$, 相反的情形可用$\mathbf{v}$的负向量替换之, 如果$\varepsilon>0$充分小, 而$\dot a\in S$, 那么, 对所有的$i=1,\cdots,m$, 都有$f_i(\dot a+\varepsilon\mathbf{v})>0$, 只要取$\varepsilon<\min_i(f_i(\dot a)/|Df_i(\mathbf{v})|)$即可. 从而, 对所有这些$\varepsilon$, 都有$\dot a+\varepsilon\mathbf{v}\in S$. 但是, $f(\dot a+\varepsilon\mathbf{v})=f(\dot a)+\varepsilon Df(\dot a)$, 所以, $f(\dot a)$不是f的一个极大值. □

现在可以证明我们的基本结果.

定理 设仿射线性函数f在多面体S上有上限. 那么, 它必在S的某些界面点上取得自己的极大值, 而这些点本身同样构成一个多面体. 如果S是有界的, 那么, f在多面体S的某些顶点上取得自己的极大值.

证明 我们对空间$\mathbb{A}$的维数用归纳法. 当$\dim\mathbb{A}$=0时, 显然. 设$\dim\mathbb{A}=n$, 且设对于维数小于n的情形, 命题已证. 设多面体S由一组不等式$f_1\geqslant 0,\cdots,f_m\geqslant 0$给出. 因

为集合S是闭的, 所以, 它的上方有界的函数f必在某个点$\dot{a}$上取极大值. 如果$f_1(\dot{a}) > 0, \cdots, f_m(\dot{a}) > 0$, 那么, 根据引理3, f只能是个常量;特别地, 它在整个S上都取得自己的极大值. 现在, 我们可以认为, 有某个i使得$f_i(\dot{a}) = 0$. 这意味着, f在非空多面体S_i的一个点上取得了极大值, 而S_i是S的一个界面而且属于$n-1$维仿射子空间$\{\dot{a}|f_i(\dot{a}) = 0\}$, 因为$f_i$不是常量. 按归纳假设, 限制在$S_i$上的$f$的极大值应该取在多面体$S_i$的某个界面上的所有点. 据引理1和引理2, 它也是原来的多面体的一个界面, 而且也是个多面体. 因为在S_i上定义它的不等式具有一个可以推广到整个空间$\mathbb{A}$上去的左端, 从而可以补充为$f_i = 0$.

现在, 按照S的仿射包络的维数进行归纳, 我们要证明, 任意有界的多面体必有顶点. 事实上, 对于零维的情形, 这是显然的. 设维数大于零. 我们可以认为, 多面体S的仿射包络就是整个$\mathbb{A}$. 在$\mathbb{A}$上任取一个非常数仿射线性函数. 它应当在S上有极大值, 因为S是有界的而且是闭的. 可见, S有非空的界面, 它由所有的取得该极大值的点组成. 它是个有界的多面体, 它的仿射包络的维数必严格地变小, 按归纳法假设, 它应该有顶点, 而按引理3, 它同样应该是多面体S的顶点.

最后, 设S是有界的, 且T是S的多面体界面, 原始函数f在它上面取得自己的极大值, 那么, T的任意顶点, 我们已经证明了是存在的, 都是多面体S的一个未知顶点. □

习　　题

1. 证明, 所有的有界的多面体都是自己的顶点集合的凸包络.

2. 利用习题1和第4章§3的定理13证明, 在有界的多面体上的线性函数必在自己的一个顶点上达到极大值.

§4 非负矩阵

1. 生产上的论据 追踪[12, 15], 我们来讲述利用相当知名的列昂季耶夫[1] 经济模型的生产计划问题. 某个联合企业有n个工厂$F_j, 1 \leqslant j \leqslant n$. 在工厂$F_j$生产$P_j$产品. 为了生产$P_k$产品需要使用$a_{jk} \geqslant 0$个单位的$P_j$产品, $j \neq k$(自然地, 设$a_{jj} = 0$). 联合企业为了生产x_k个P_k产品, $k = 1, 2, \cdots, n$就总共需要$\sum\limits_{k=1}^{n} a_{jk}x_k$个单位的$P_j$. 这样一来, 给市场剩下的是

$$y_j = x_j - \sum_{k=1}^{n} a_{jk}x_k \tag{1}$$

个单位的P_j产品.

计划问题可采用如下形式: 对于给定的市场需求$\mathbf{y} = (y_1, \cdots, y_n) \in \mathbb{R}^n$, 且$y_j \geqslant 0$, 需要找出生产向量$\mathbf{x} = (x_1, \cdots, x_n)$, 且$x_j \geqslant 0$, 满足条件(1). 设$A = (a_{jk})$, 我

1) 瓦西里 · 瓦西里耶维奇 · 列昂季耶夫(1905—1999)是圣彼得堡大学毕业生, 随后为加鲁阿大学教授, 著名经济学家, 诺贝尔奖获得者.

们把(1)改写成矩阵形式

$$\mathbf{y} = (E - A)\mathbf{x}. \tag{1'}$$

现在, 我们对于系数为非负实数的向量和矩阵用符号

$$\mathbf{x} \geqslant 0,\ \mathbf{y} \geqslant 0,\ A \geqslant 0 \tag{2}$$

表示, 并简称是**非负向量**, **非负矩阵**. 在严格不等式的情形, 就说是**正向量**和**正矩阵**(不要与正定矩阵混淆).

现在, 运用§1的第5目中矩阵谱半径的初等结论(或者把它与它在$\mathbb{R}^n$上的线性算子等同起来). 如果$r(A) < 1$, 那么, 作为§1例3的推论, 矩阵$E - A$是可逆的, 而且$(E - A)^{-1} = \sum_{k\geqslant 0} A^k$. 因为, 按定义, 矩阵$A$是非负的, 所以它的任意次幂$A^k$都是非负的, 在这种情形下, 矩阵$(E - A)^{-1}$必然也是非负的. 从而, 矩阵(1′)对应的所需要的解$\mathbf{x} \geqslant 0$可由公式

$$\mathbf{x} = (E - A)^{-1}\mathbf{y}.$$

给出. 经济学技术领域未必关注条件$r(A) < 1$. 但是, 利用§1的第5目的不等式3): $r(A) \leqslant \max_k \sum_{j=1}^n a_{jk}$, 我们可以靠近它. 于是, $\sum_{j=1}^n a_{jk} < 1, k = 1, 2, \cdots, n$就是解决这个问题的一个充分条件. 这个条件允许经济学上的关注. 实际上, $\sum_{j=1}^n a_{jk}$是制造一个单位P_k产品时工厂F_k要支付的成本. 从而, 要求$\sum_{j=1}^n a_{jk} < 1$表示工厂F_k的工作是有赢余的.

这样一来, 我们有

定理1 *如果每个工厂的生产都是有赢余的, 那么, 计划问题是可解的, 而且只有一种解决方式.*

2. 非负矩阵的性质 按照定理1, 具有满足条件(2)的矩阵$A \geqslant 0$的方程组(1), 对任意$\mathbf{y} \geqslant 0$, 都有唯一解, 也就是$\det(E - A) \neq 0$. 现在, 要注意, 如果$A \geqslant 0$且$0 \leqslant \lambda \leqslant 1$, 那么, 矩阵$\lambda A \geqslant 0$同样也满足条件(2), 所以,

$$\det(E - \lambda A) \neq 0, 0 \leqslant \lambda \leqslant 1.$$

行列式$\det(E - \lambda A)$在$\lambda = 0$时是正的, 而且是对λ连续的, 所以, 当$\lambda = 1$时, 它也是正的. 也就是说, 有

定理2 *设*$A = (a_{jk}) \geqslant 0$ *且*$\sum_{j=1}^n a_{jk} < 1$ *对任意*$k = 1, 2, \cdots, n$ *都成立, 那么,* $\det(E - A) > 0$.

非负矩阵是博弈论、组合理论、最优化问题、数理经济(线性规划和动态规划)、概率论等分支研究以及遗传学中重要的离不开的工具.

设P是一个由置换$\pi \in S_n$对应的矩阵(特别地, 可参见[BA Ⅰ]第2章§3的习题6.)例如, 当n=3时且$\pi = (1\ \ 2\ \ 3)$时, 我们有

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

显然, ${}^tP = P^{-1}$. 相似变换$A \mapsto P^{-1}AP$同时对矩阵$A \in M_n(\mathbb{R})$的行和列施以变换.

定义1 设, 当$n > 1$时, 能够找到一个矩阵P, 使得

$$P^{-1}AP = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

其中A_{11}和A_{22}都是阶数小于n的方阵. 那么, 称A是个**可约化矩阵**. 如果这样的矩阵P不存在, 则称A是个**不可约化矩阵**.

显然, 矩阵$A > 0$是不可约化的, 因为不管对A的行或列进行置换, 它的角上都不会出现零.

对于非负矩阵, 一个基础性的成果是下面的佩龙-弗罗贝尼乌斯(1907—1912)经典定理, 它后来被威兰多改进了.

定理3 设$A = (a_{ij}) \in M_n(\mathbb{R})$是任意一个非负的不可约化矩阵, 它的特征根是$\lambda_0, \lambda_1, \cdots, \lambda_{n-1}$. 那么

i) A有一个正的特征值$r = r(A)$, 它的代数重数是1;

ii) 特征值r对应正的特征向量$\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}$;

iii) 如果$\{\lambda_0, \lambda_1, \cdots, \lambda_{k-1}\}$是所有的模等于$r$的特征根的集合, 那么$\lambda_i = \theta_j r, 0 \leqslant j \leqslant k-1$, 其中$\theta_j = e^{2\pi ij/k}$, 而且在$\mathbb{C}$中有$n$个点$\lambda_s, 0 \leqslant s \leqslant n-1$, 它们对于绕坐标原点的多重$2\pi/k$角旋转是不变的;

iv) 如果$k > 1$, 那么, 可以找到置换矩阵P, 使得

$$P^{-1}AP = \begin{pmatrix} 0 & A_{12} & 0 & \cdots & 0 \\ 0 & 0 & A_{23} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & A_{k-1,k} \\ A_{k1} & 0 & 0 & \cdots & 0 \end{pmatrix},$$

其中$A_{j,j+1}$是$n_j \times n_{j+1}$阶矩阵而$A_{k,1}$是$n_k \times n_1$阶矩阵;

v) 如果至少有1个j使$a_{jj} > 0$, 那么, $k = 1$;

vi) 如果能找到$i \neq j$使得$a_{ij}a_{ji} > 0$, 那么, $k \leqslant 2$.

定理3的证明相当长, 在这里就不加以推导了.

3. 随机矩阵 我们回顾(见[BA Ⅰ])第2章§3的习题4)下面的

定义2 称矩阵$P=(p_{ij})\in M_n(\mathbb{R})$是**随机的**, 如果

$$P\geqslant 0,\sum_{j=1}^{n}p_{ij}=1,\quad i=1,2,\cdots,n.$$

此外, 如果$\sum\limits_{i=1}^{n}p_{ij}=1,\quad j=1,2,\cdots,n$, 那么, 称矩阵$P$是**双随机的**.

置换矩阵就是一种特殊类型的双随机矩阵.

定理4 非负矩阵P具有随机性质, 当且仅当, $P\mathbf{e}=\mathbf{e},\mathbf{e}=[1,1,\cdots,1]$. 此外, 对每个随机矩阵$P$都有$r(P)=1$.

证明 如果矩阵$P=(p_{ij})$是随机的, 那么, 显然$P\mathbf{e}=\mathbf{e}$, 所以1是P的一个特征值. 反之, 性质$P\mathbf{e}=\mathbf{e}$只是P的随机性的另一个记法而已.

其次, 对于随机矩阵P, 由$P\mathbf{x}=\lambda\mathbf{x},\mathbf{x}=[x_1,\cdots,x_n]$可得$\sum\limits_{j=1}^{n}p_{ij}x_j=\lambda x_i$, 所以, 只要设$|x_m|=\max\limits_{1\leqslant j\leqslant n}\{x_j\}$, 我们即可推出不等式

$$|\lambda|\cdot|x_i|\leqslant\sum_{j=1}^{n}p_{ij}|x_j|\leqslant|x_m|\sum_{j=1}^{n}p_{ij}=|x_m|.$$

特别地, 当$i=m$时, 我们就有

$$|\lambda|\cdot|x_m|\leqslant|x_m|\Rightarrow|\lambda|\leqslant 1\Rightarrow r(P)=1.$$ □

下面的定理建立了随机矩阵与一般非负矩阵的联系.

定理5 设A是任意一个非负矩阵, 正的向量$\mathbf{c}=[c_1,\cdots,c_n]$是$A$的属于特征值$r=r(A)$的一个特征向量. 再设$C=\mathrm{diag}(c_1,\cdots,c_n)$. 那么, 矩阵

$$P=(p_{ij})=\frac{1}{r}C^{-1}AC$$

必然是随机的.

证明 根据条件, $A\mathbf{c}=\gamma\mathbf{c}$, 它等价于一组等式

$$\sum_{j=1}^{n}a_{ij}c_j=rc_i,\quad i=1,2,\cdots,n.$$

因为$p_{ij}=r^{-1}c_i^{-1}a_{ij}c_j$, 所以

$$\sum_{j=1}^{n}p_{ij}=r^{-1}c_i^{-1}\sum_{j=1}^{n}a_{ij}c_j=r^{-1}c_i^{-1}\cdot rc_i=1.$$ □

容易看出来, 矩阵$P=(p_{ij}),Q=(q_{ij})\in M_n(\mathbb{R})$的随机性可导出它们的乘积$PQ=R=(r_{ij})$的随机性:

$$\sum_{j=1}^{n}r_{ij}=\sum_{j=1}^{n}\sum_{k=1}^{n}p_{ik}q_{kj}=\sum_{k=1}^{n}p_{ik}\sum_{j=1}^{n}q_{kj}=\sum_{k=1}^{n}p_{ik}=1.$$

特别地, 任意随机矩阵的乘幂都必然还是随机矩阵.

我们在这里没有可能去证明有趣的**遍历性**, 即极限

$$\tilde{P} = \lim_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} P^j$$

的存在性, 其中$\tilde{P}$仍然是一个随机矩阵且$\tilde{P}^2 = \tilde{P} = \tilde{P}P = P\tilde{P}$. 如果$P$是个不可约化的随机矩阵, 而且有

$$\lambda \in \mathrm{Spec}(P), |\lambda| = 1 \Rightarrow \lambda = 1,$$

那么, 同样存在极限

$$P^{\infty} = \lim_{k\to\infty} P^k. \tag{3}$$

最后, 如果P是双随机的且$\dim\ \mathrm{Ker}(P-E) = 1$. 那么,

$$\tilde{P} = \lim_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} P^j = \begin{pmatrix} 1/n & 1/n & \cdots & 1/n \\ \cdots & \cdots & \cdots & \cdots \\ 1/n & 1/n & \cdots & 1/n \end{pmatrix}.$$

任意一个随机矩阵P, 显然地, 都把一个概率向量$\mathbf{x} = [x_1, \cdots, x_n], x_j \geqslant 0$, $\sum_{i=1}^{n} x_i = 1$仍然变成一个概率向量. 而且, 在$P > 0$的情形, 任意一个概率向量$\mathbf{x}$都对应一个正的向量$P\mathbf{x}$. 也就是说, 按照这个道理, 随机矩阵在概率论中起极重要的作用。设给定了某个物理系统$\mathbf{S}$, 它可以处于$s_1, \cdots, s_n$中的某一种状态. 设系统在离散时刻$t_0 < t_1 < t_2 < \cdots$观测, 而且$P_k = [p_1(t_k), \cdots, p_n(t_k)]$是随机向量. 其中, $p_j(t_k)$是当$j = 1, 2, \cdots, n; k = 0, 1, 2, \cdots$时, 系统在$t_k$时刻处于状态$s_j$的绝对概率. 现在, 再设(这个假设是完全可以实现的), 由t_{k-1}时刻系统处于s_j状态过渡到t_k时刻系统$\mathbf{S}$处于s_i状态的已知的条件概率是P_{ij}. 按照概率论规则

$$p_i(t_k) = \sum_{j=1}^{n} p_{ij} p_j(t_{k-1}).$$

令$P = (p_{ij})$, 把此式改写成矩阵形状: $\mathbf{p}_k = P\mathbf{p}_{k-1}$. 最后, 如果系统$\mathbf{S}$的初始状态是$s_r$, 那么$\mathbf{p}_0$是这样一个列向量: 它的第$r$个位置是1, 其余位置是零. 于是

$$\mathbf{p}_k = P^k \mathbf{p}_0.$$

写在这里的过程可称之为以P为过渡矩阵的**齐次的马尔可夫**[1]**链**(通常的马尔可夫链, 条件$p_{ij} = p_{ij}(t_k; t_{k-1})$由时刻$t_k$决定). 很自然地, 要关心序列$\{\mathbf{p}_k\}$的极限状态, 也就是实际上的极限(3).

1) 安德烈 · 安德烈耶维奇 · 马尔可夫(1856—1922)是杰出的俄罗斯数学家, 彼得堡数学学派最著名的代表之一.

例　过渡矩阵

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

对应的基本状态形式“是”, “否”, 且在这个过程畸变中转换“是→否”的概率是p, 而转换“否→是”的概率是q; $1-p$就是“是→是”的概率, $1-q$就是“否→否”的概率.

记$P = Q + E$, 其中

$$Q = \begin{pmatrix} -p & p \\ q & -q \end{pmatrix},$$

我们就有$Q^2 = -(p+q)Q$. 在这里, $0 \leqslant p, q \leqslant 1$, 所以$p+q=0 \Rightarrow P = E$, 即为平凡情形. 故可设$p+q>0$, 从而,

$$\begin{aligned} P^k = (Q+E)^k &= \sum_{j=0}^{k} \binom{k}{j} Q^k = E + \sum_{j=1}^{k} (-1)^{j-1}(p+q)^{j-1} Q \\ &= E + \frac{1}{p+q} Q - \frac{(1-p-q)^k}{p+q} Q. \end{aligned}$$

显然, $-1 \leqslant 1-p-q \leqslant 1$. 当$1-p-q=1$时, 我们再次回到平凡的情形$P=E$. 而当$1-p-q=-1$, 即$p=q=1$时, 因为

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad P^{2k} = E, \quad P^{2k+1} = P,$$

所以. 极限P^∞是不存在的.

现在, 研究一般情形, $-1 < 1-p-q < 1$. 我们有$\lim\limits_{k\to\infty}(1-p-q)^k = 0$, 所以,

$$P^\infty = \lim_{k\to\infty} p^k = E + \frac{1}{p+q} Q = \begin{pmatrix} \dfrac{q}{(p+q)} & \dfrac{p}{(p+q)} \\ \dfrac{q}{(p+q)} & \dfrac{p}{(p+q)} \end{pmatrix}.$$

特别地, 当$0 < p = q < 1$时, 我们有

$$\lim_{k\to\infty} P^k = \begin{pmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}.$$

作为结尾, 我们应注意. n阶的半魔幻矩阵的空间$\mathrm{SMag}_n(\mathbb{Q})$的所有非负矩阵$A = (a_{ij})$(见第1章§1的例8)一定可以化成双随机矩阵$P = (p_{ij})$. 事实上, $\sum\limits_j a_{ij} = \alpha = \sum\limits_i a_{ij}, 1 \leqslant i, j \leqslant n, 0 \neq \alpha \in \mathbb{Q}$, 那么, 只要令$p_{ij} = 1/\alpha$就够了. 这可以算是定理5的一个补充说明.

§5 罗巴切夫斯基几何

1. 罗巴切夫斯基空间 设V是实数域$\mathbb{R}$上的一个$n+1$维的向量空间，$\mathbb{P}(V)$是V生成的射影空间. 设q是V上的一个非退化的二次型，它的惯性指标是n，而负惯性指数为1;f是与q配极的对称双线性型. 正如我们在第一章就已经知道的，必存在一个基底，在这个基底之下，q取标准形式

$$q(\mathbf{x}) = -x_0^2 + x_1^2 + x_2^2 + \cdots + x_n^2. \tag{1}$$

我们可以认为对(V, q)是一个"伪欧几里得"空间(闵可夫斯基空间)，而现在，我们从另一个方面靠近这个目标则更为有益.

设

$$\tilde{\mathbf{x}} \in \mathbb{P}(V), \quad \mathbf{x} = x_0\mathbf{e}_0 + \cdots + x_n\mathbf{e}_n,$$

则可把$x_0, x_1, \cdots, x_n$看成是这一个点在给定的射影坐标系之下的齐次坐标. 用方程式$q(\mathbf{x}) = 0$给定射影二次曲面(在V中，我们称它是个锥体).

定义1 由条件$q(\mathbf{x}) < 0$，即

$$x_1^2 + x_2^2 + \cdots + x_n^2 < x_0^2 \tag{2}$$

给定的集合$\Lambda \subset V$(在相对论中的光锥)是锥体的"**内部**". 所有的经过坐标原点且处于锥体内部的直线对应的集合

$$\tilde{\Lambda} = P_r(\Lambda) \subset \mathbb{P}(V),$$

就称为是n维**罗巴切夫斯基空间**.

事实上，$\tilde{\Lambda}$整个地包含在由条件$x_0 \neq 0$决定的仿射图$(\mathbb{E}_0, \Phi_0)$之中. 这个条件对于满足条件(2)是必要的. 在仿射坐标系之下

$$y_i = x_i/x_0, i = 1, 2, \cdots, n,$$

空间$\tilde{\Lambda}$由不等式

$$y_1^2 + y_2^2 + \cdots + y_n^2 < 1 \tag{2'}$$

给出. 也就是说，$\Phi_0(\tilde{\Lambda})$可以画成形如以点$\dot{e}_0$为中心，以1为半径的$n-1$维的一个开球(图24).

现在，我们回顾一下，$\mathbb{E}_0 = \mathbf{e}_0 + V_0$，其中$\mathbf{e}_0$是这个基底的第一个向量，在这个基底之下，$q(\mathbf{x})$形式如(1)，而$V_0$是由条件$x_0 = 0$决定的向量的超平面. 换言之，$\mathbb{E}_0$是与向量空间$V_0$相伴的仿射空间. 在$V_0$上，二次型$q_0 = q|_{V_0}$完全确定，它在基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下形如$q_0(\mathbf{x}) = x_1^2 + \cdots + x_n^2$，并且可由此推出它是正定的. 在$\mathbb{E}_0$上借助$q_0$可以引进欧几里得距离函数. 这样一来，$\mathbb{E}_0$就有了欧几里得空间结构.

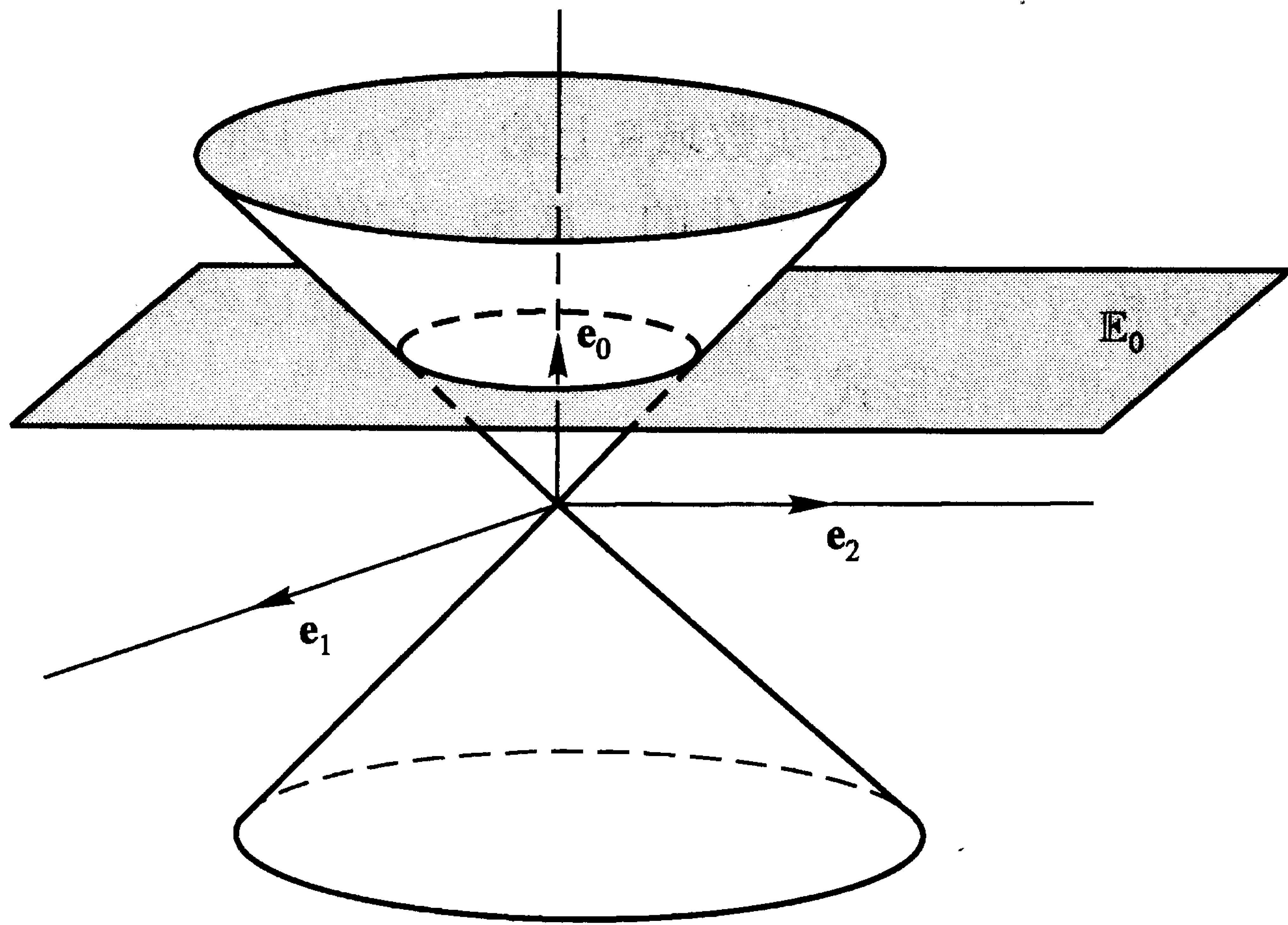


图24

现在，我们来推广已有的结论并尝试建立与空间$\mathbb{P}(V)$的仿射图$(\mathbb{E}_\mathbf{e}, \Phi_\mathbf{e})$的每个点$\tilde{\mathbf{e}} \in \tilde{\Lambda}$的联系. 因为$q(\mathbf{e}) < 0$且$\widetilde{\lambda \mathbf{e}} = \tilde{\mathbf{e}}$, 所以，可以认为，向量e满足正规条件$q(\mathbf{e}) = -1$. 令

$$V_\mathbf{e} = \{\mathbf{x} \in V \,|\, f(\mathbf{e}, \mathbf{x}) = 0\}, \tag{3}$$

其中f是与q配极的双线性型. 可见，$V_\mathbf{e}$是e的伪正交补. 进而，再设

$$\mathbb{E}_\mathbf{e} = \mathbf{e} + V_\mathbf{e}.$$

二次型$q_\mathbf{e} = q\,|_{V_\mathbf{e}}$是正定的. 实际上，把向量$\mathbf{x} \in V$记成$\mathbf{x} = \alpha \mathbf{e} + \mathbf{x}'$, $\mathbf{x} \in V_\mathbf{e}$, 并再次利用定义式(3), 我们可以看出来

$$q(\mathbf{x}) = -\alpha^2 + q(\mathbf{x}') = -\alpha^2 + q_\mathbf{e}(\mathbf{x}').$$

由此可见，二次型$q_\mathbf{e}$的负惯性指标应该等于零.

这就是说，$(V_\mathbf{e}, q_\mathbf{e})$是个带有纯量乘积$(\mathbf{x}' \,|\mathbf{y}')_\mathbf{e}$的欧几里得向量空间. 这就允许把$\mathbb{E}_\mathbf{e}$看成是一个(点的)欧几里得空间. 按通常方式定义了映射

$$\Phi_\mathbf{e} : \mathbb{P}(V) \backslash \mathbb{P}(V_\mathbf{e}) \to \mathbb{E}_\mathbf{e},$$

$\Phi_\mathbf{e}(\tilde{\mathbf{x}})$是直线$\langle \mathbf{x} \rangle$与$\mathbb{E}_\mathbf{e}$的交点. 特别地，

$$\Phi_\mathbf{e}(\tilde{\Lambda}) = \Lambda \cap \mathbb{E}_\mathbf{e}.$$

向量$\mathbf{x} = \mathbf{e} + \mathbf{x}'$, $\mathbf{x}' \in V_\mathbf{e}$属于集合$\Lambda$, 当且仅当，$q(\mathbf{x}) = q_\mathbf{e}(\mathbf{x}') - 1 < 0$, 即$q_\mathbf{e}(\mathbf{x}') < 1$. 而这就意味着，在空间$\tilde{\Lambda}$的仿射图$\langle \mathbb{E}_\mathbf{e}, \Phi_\mathbf{e} \rangle$上，可画成形如以点$\dot{e} \in \Phi_\mathbf{e}(\tilde{\mathbf{e}})$为中心，以1为半径的开的欧几里得球.

2. 罗巴切夫斯基空间的运动 伪正交群$O(q) = O(n,1) \subset GL(V)$, 或者, 同样的也就是二次型$q$的自同构群, 由如下的满足条件(4)的线性算子$\mathcal{A} \in GL(V)$组成,

$$q(\mathcal{A}\mathbf{x}) = q(\mathbf{x}), \quad \forall \mathbf{x} \in V. \tag{4}$$

将(2)和(4)加以比较, 我们可以看出, $\mathcal{A}(\Lambda) = \Lambda$. 研究在满射$\pi : GL(V) \to PGL(V)$之下群$O(q)$的像$\widetilde{O(q)}$(见第5章, §3, 第6目)

$$\widetilde{O(q)} = \left\{ \tilde{\mathcal{A}} = \pi(\mathcal{A}) \,|\, \mathcal{A} \in O(q) \right\}.$$

因为$\mathcal{A}(\Lambda) = \Lambda$且$\tilde{\mathcal{A}} \cdot \tilde{\mathbf{x}} = \widetilde{\mathcal{A}\mathbf{x}}$, 所以

$$\tilde{\mathbf{x}} \in \tilde{\Lambda} \Leftrightarrow \mathbf{x} \in \Lambda \Rightarrow \mathcal{A}\mathbf{x} \in \Lambda \Rightarrow \tilde{\mathcal{A}} \cdot \tilde{\mathbf{x}} \in \tilde{\Lambda}.$$

也就是说, $\widetilde{O(q)}$是射影群$PGL(V)$的一个包含$\tilde{\Lambda}$的子群.

定义2 *称群$\widetilde{O(q)}$的元素是一个***运动**, *而群$\widetilde{O(q)}$本身是一个m维罗巴切夫斯基空间$\tilde{\Lambda}$的***运动群**. *作用在$\tilde{\Lambda}$上的群$\widetilde{O(q)}$ 对应的几何学, 即称为***双曲几何学***或者***罗巴切夫斯基几何**[1)]

在这个定义中采用的术语"运动"可能令人不太习惯,. 通常, 把它和集合上保持某种集合度量不变的变换联想到一起. 我们的基本问题是在$\tilde{\Lambda}$上导出一个度量, 使得群$\widetilde{O(q)}$名副其实. 但是, 眼前, 我们先停下来研究群$\widetilde{O(q)}$的另外一个性质, 它对于几何学的含义丰富的定义同样是必要的.

定理1 *群$\widetilde{O(q)}$作用在$\tilde{\Lambda}$上是可迁的, 也就是说, 在罗巴切夫斯基几何学中, 所有的点都是$\widetilde{O(q)}$同余的.*

证明 设$\tilde{\mathbf{c}}$, $\tilde{\mathbf{e}}$是空间$\tilde{\Lambda}$的任意两个点, 正如我们已经看到的, 不失一般性, 可以认为, $q(\mathbf{c}) = q(\mathbf{e}) = -1$. 在向量超平面$V_{\mathbf{c}}$和$V_{\mathbf{e}}$中(见(3)), 我们选取标准正交基底$(\mathbf{c}_i')$, $(\mathbf{e}_i')$, $1 \leqslant i \leqslant n$ (分别相对于纯量乘积$(*|*)\mathbf{c}$ 和$(*|*)\mathbf{e}$标准正交). 与q配极的双线性型f在空间V的每个基底$(\mathbf{c}, \mathbf{c}_i')$, $(\mathbf{e}, \mathbf{e}_i')$之下的矩阵$F$都有同一种形式$\mathrm{diag}(-1, 1, 1, \cdots, 1)$. 所以, 由等式$\mathcal{A}\mathbf{c} = \mathbf{e}, \mathcal{A}\mathbf{c}_i' = \mathbf{e}_i', 1 \leqslant i \leqslant n$, 给定的线性算子, 显然是二次型$q$ (或者f)的自同构, 也就是$\mathcal{A} \in O(q)$. 同时, $\tilde{\mathcal{A}} \cdot \tilde{\mathbf{c}} = \widetilde{\mathcal{A}\mathbf{c}} = \tilde{\mathbf{e}}$, 所以, $\tilde{\mathcal{A}}$是罗巴切夫斯基空间中把点$\tilde{\mathbf{c}}$变成点$\tilde{\mathbf{e}}$的一个运动.

再给出关于群$\widetilde{O(q)}$的某些信息.

定理2 *任意点$\tilde{\mathbf{e}} \in \tilde{\Lambda}$的稳定子群$\widetilde{O(q)}_{\mathbf{e}}$. 同构于在仿射图$(\mathbb{E}, \Phi_{\mathbf{e}})$上绕点$\dot{e} = \Phi_{\mathbf{e}}(\tilde{\mathbf{e}})$的旋转群$SO(n)$.*

证明 我们的目的是指明, 罗巴切夫斯基空间$\tilde{\Lambda}$中, 在点$\tilde{\mathbf{e}} \in \tilde{\Lambda}$处不动的运动$\tilde{\mathcal{A}}$必然对应$\mathbb{E}_{\mathbf{e}}$中绕$\dot{e}$点的一个旋转, 而且, 反之亦然.

1) 为尊崇伟大的俄罗斯几何学家尼古拉 · 伊万诺维奇 · 罗巴切夫斯基(1792—1856). 他揭示并第一个讲述了几何基础(1829年). 他的著作《关于几何基础》及其他后继关于罗巴切夫斯基平面模型的工作中, 已被证明了是独立于欧几里得五公设的.

这就是说, 设$\tilde{\mathcal{A}} \cdot \tilde{\mathbf{e}} = \tilde{\mathbf{e}}$, 即$\mathcal{A}\mathbf{e} = \lambda\mathbf{e}$. 根据条件$q(\mathbf{e}) = -1$. 因为$\mathcal{A} \in O(q)$, 所以,

$$q(\lambda\mathbf{e}) = q(\mathbf{e}) \Rightarrow \lambda = \pm 1.$$

不失一般性, 我们可以认为$\lambda = 1$. 如有必要时可以用$-\mathcal{A}$来代替$\mathcal{A}$. 由(3)可得$\mathcal{A}(V_{\mathbf{e}}) = V_{\mathbf{e}}$. 设$\mathcal{A}_{\dot{e}} = \mathcal{A} \mid \mathbb{E}_{\mathbf{e}}$,$\mathcal{A}'_{\mathbf{e}} = \mathcal{A}_{\dot{e}}|_{V_{\mathbf{e}}}$, 我们有

$$\mathcal{A}_{\dot{e}}(\mathbf{e} + \mathbf{x}') = \mathbf{e} + \mathcal{A}'_{\mathbf{e}}\mathbf{x}' \tag{5}$$

对任意点$\dot{x} := \mathbf{x} = \dot{\mathbf{e}} + \mathbf{x}' \in \mathbb{E}_{\mathbf{e}}, \mathbf{x}' \in V_{\mathbf{e}}$都成立.

这样一来, $\mathcal{A}'_{\mathbf{e}}$就是$V_{\mathbf{e}}$上的一个正交线性变换, 而$\mathcal{A}_{\dot{e}}$是空间$\mathbb{E}_{\mathbf{e}}$的绕点$\dot{e} := \mathbf{e}$的一个旋转. 因为

$$\tilde{\mathbf{x}} \in \tilde{\Lambda} \Rightarrow \mathbf{x} = \mu(\mathbf{e} + \mathbf{x}'), \mu \neq 0,$$

所以, 考虑到(5), 我们就有

$$\begin{aligned} \Phi_{\mathbf{e}}(\tilde{\mathcal{A}} \cdot \tilde{\mathbf{x}}) &= \Phi_{\mathbf{e}}(\widetilde{\mathcal{A}\mathbf{x}}) = \Phi_{\mathbf{e}}(\mu(\dot{e} + \mathcal{A}'_{\mathbf{e}}\mathbf{x}')) \\ &= \dot{e} + \mathcal{A}'_{\mathbf{e}}\mathbf{x}' = \mathcal{A}_{\dot{e}}(\mathbf{e} + \mathbf{x}') = \mathcal{A}\,(\Phi_{\mathbf{e}}\tilde{\mathbf{x}}), \end{aligned}$$

就是说, 旋转$\mathcal{A}_{\dot{e}}$就表达了$\tilde{\mathcal{A}} : \tilde{\Lambda} \to \tilde{\Lambda}$在仿射图$(\mathbb{E}_{\mathbf{e}}, \Phi_{\mathbf{e}})$的作用.

照样, 如果$\mathcal{A}_{\dot{e}}$是空间$\mathbb{E}_{\mathbf{e}}$绕点$\dot{e}$的旋转, 而它的线性部分为$\mathcal{A}'_{\mathbf{e}}$, 那么, 由等式$\mathcal{A}\mathbf{e} = \mathbf{e}, \mathcal{A}\mathbf{x}' = \mathcal{A}'_{\mathbf{e}}\mathbf{x}'$决定的线性算子$\mathcal{A} : V \to V$, 显然, 是保持仿射超平面$\mathbb{E}_{\mathbf{e}} \subset V$不变的二次型$q$的一个自同构, 同时, $\tilde{\mathcal{A}} \in \widetilde{O(q)}$作用在$\tilde{\Lambda}$上刚好与$\mathcal{A}_{\dot{e}}$作用在$\mathbb{E}_{\mathbf{e}}$上相对应. □

3. 罗巴切夫斯基度量　以下我们在欧几里得空间$\mathbb{E}_0 = \mathbf{e}_0 + V_0$中, 将开球$(2')$的中心点$\dot{e}_0$固定, 并且返回到仿射图$(\mathbb{E}_0, \Phi_0)$上来.

引理1　设$\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2 \in \tilde{\Lambda}$且$\tilde{\mathbf{a}}_1 \neq \tilde{\mathbf{a}}_2$. 我们经过$\Phi_0(\tilde{\mathbf{a}}_1)$和$\Phi_0(\tilde{\mathbf{a}}_2)$引一条仿射空间$\mathbb{E}_0$的直线$L$.

那么, L与球面

$$S^{n-1} : y_1^2 + y_2^2 + \cdots + y_n^2 = 1 \tag{6}$$

相交于两个不同的点.

证明　直线L上的点可以以$\dot{p} + t\mathbf{a}'$形式给出, 其中$\dot{p}$是$\mathbb{E}_0$的某个点, $\mathbf{0} \neq \mathbf{a}'$是$V_0$的一个向量, 而$t$是任意实数(在直线上点的坐标). 如果, $\mathbf{a}' = \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n, \dot{p} = (\beta_1, \cdots, \beta_n)$,那么, 点$\dot{r} = \dot{p} + t\mathbf{a}'$的仿射坐标形如$y_i = \beta_i + t\alpha_i$. 直线与球面(6)的交点可由方程

$$\alpha t^2 + \beta t + \gamma = \sum_{i=1}^{n} (t\alpha_i + \beta_i)^2 - 1 = 0$$

得到. 二次三项式$\alpha t^2 + \beta t + \gamma$在$t^2$项带有正系数$\alpha = \sum_i \alpha_i^2$, 在直线$L$上的点$\Phi_0(\tilde{\mathbf{a}}_1)$, $\Phi_0(\tilde{\mathbf{a}}_2)$处取负值, 这是因为, 对它们而言, $\sum_i y_i^2 - 1 < 0$. 这就意味着$\alpha t^2 + \beta t + \gamma$有两个不同的实根. □

设$\Phi_0(\tilde{\mathbf{b}}_1)$, $\Phi_0(\tilde{\mathbf{b}}_2)$是$L$和球面(6)的交点, 选取时按这样的顺序, 即, 使得在直线上点$\Phi_0(\tilde{\mathbf{b}}_1)$, $\Phi_0(\tilde{\mathbf{a}}_1)$, $\Phi_0(\tilde{\mathbf{a}}_2)$, $\Phi_0(\tilde{\mathbf{b}}_2)$的坐标$t_k$按递增的方式(如有需要, 可以把$\tilde{\mathbf{a}}_1$和$\tilde{\mathbf{a}}_2$的位置互换). 我们得到

$$\Delta(\dot{a}_1, \dot{a}_2) := \delta \cdot |\dot{a}_1\dot{a}_2| = \left[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_1\right], \tag{7}$$

其中, 按惯例. $\left[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\right]$是四点的重比(见第5章§3). 我们有$\tilde{\mathbf{a}}_1 \neq \tilde{\mathbf{a}}_2$, 但是, 关系式(7)当$\tilde{\mathbf{a}}_1 = \tilde{\mathbf{a}}_2$的时候亦然保持, 只要$\delta \cdot |a_1a_2|=1$且$L$是任意的经过$\Phi_0(\tilde{\mathbf{a}}_1)$的直线即可.

引理2　*如果$\tilde{a}_1 \neq \tilde{a}_2$, 那么, $\Delta(\dot{a}_1, \dot{a}_2) > 1$.*

证明　可以把经过点$\Phi_0(\tilde{\mathbf{a}}_1)$, $\Phi_0(\tilde{\mathbf{a}}_2)$的直线$L$看成是射影直线$\mathbb{P}'$的仿射图. 我们在$\mathbb{P}'$上定义另外一个仿射图, 把点$\tilde{\mathbf{b}}_1$作为无穷远点. 在这个仿射图中, 我们选取一个坐标系, 把点$\tilde{\mathbf{b}}_2$取作零点, 而把$\tilde{\mathbf{a}}_2$取作单位(显然, 这意味着, 它是第5章中重比性质的记法中最好的), 于是, 依据第5章§3的公式(21), 我们得到$\left[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\right] = x$是$\tilde{\mathbf{a}}_1$的仿射坐标. 因为$\Phi_0(\tilde{\mathbf{a}}_1)$位于$\Phi_0(\tilde{\mathbf{b}}_1)$和$\Phi_0(\tilde{\mathbf{a}}_2)$之间:

$$\begin{array}{cccc} \infty & x & 1 & 0 \\ \circ & \circ & \circ & \circ \\ \Phi_0(\tilde{\mathbf{b}}_1) & \Phi_0(\tilde{\mathbf{a}}_1) & \Phi_0(\tilde{\mathbf{a}}_2) & \Phi_0(\tilde{\mathbf{b}}_2), \end{array}$$

所以, $x > 1$. 剩下来, 只要补充说一句, 重比与仿射图的选择无关. □

说明　由重比的性质, 我们有

$$\left[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\right] = \left[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_1\right]^{-1} < 1.$$

定理3　*在公式(7)中定义的表达式$\Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2)$相对于罗巴切夫斯基空间的运动是不变的, 即*

$$\Delta(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2) = \Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) \quad \forall \tilde{\mathcal{A}} \in \widetilde{O(q)}.$$

证明　设$\tilde{\mathcal{A}} \in \widetilde{O(q)}$. 如果$\tilde{\mathbf{a}}_1 = \tilde{\mathbf{a}}_2$, 那么, $\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1 = \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2$, 从而$\Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) = \Delta(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2)$.

设$\tilde{\mathbf{a}}_1 \neq \tilde{\mathbf{a}}_2$. 因为$\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2$位于空间$\mathbb{P}(V) \supset \tilde{\Lambda}$的同一条直线$\mathbb{P}'$上, 所以, $\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1$, $\tilde{\mathcal{A}}\tilde{\mathbf{a}}_2$, $\tilde{\mathcal{A}}\tilde{\mathbf{b}}_2$, $\tilde{\mathcal{A}}\tilde{\mathbf{b}}_1$位于同一条直线$\tilde{\mathcal{A}}(\mathbb{P}')$上. 其次, $\Phi_0(\tilde{\mathbf{b}}_1)$, $\Phi_0(\tilde{\mathbf{b}}_2) \in S^{n-1}$, 其中$S^{n-1}$是方程式(6)决定的球面. 按照原有的意义, $\mathcal{A}$把S^{n-1}变成自己, 因此, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{b}}_1)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{b}}_2)$就是通过$\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1)$和$\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_2)$的直线与球面$S^{n-1}$的交点.

依据第5章§3的定理4, 得到

$$\left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_1\right] = \left[\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{b}}_2, \tilde{\mathbf{b}}_1\right] = \Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2). \tag{8}$$

剩下来的是要指明$\left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_2\right] = \Delta\left(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2\right)$. 而要确信这一点, 只要把点$\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{b}}_1)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_2)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{b}}_2)$按坐标的递增方式排列在连接它们的仿射

直线上. 现在假设, 如果不是这样, 我们可以认定, 点$\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{b}}_2)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_2)$, $\Phi_0(\tilde{\mathcal{A}}\tilde{\mathbf{b}}_1)$按坐标递增顺序排列起来了. 但是, 根据引理2, 就应该有

$$\left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_2\right] = \Delta(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2) > 1,$$

从而

$$\left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_1\right] = \left[\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1\tilde{\mathcal{A}}\tilde{\mathbf{a}}_2, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{b}}_2\right]^{-1} < 1.$$

得到的不等式显然与关系式(8)和引理2矛盾. □

定义3 把量

$$\rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) = \log \Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) \tag{9}$$

称为点$\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2 \in \tilde{\Lambda}$的**罗巴切夫斯基距离**, 其中$\Delta$是由关系式(7)给出的函数.

对数的底仅仅会影响比例, 就不需指明了. 由我们在前面已经证明了的关于Δ的事实, 可以导出如下的**罗巴切夫斯基度量**ρ的性质:

i) $\rho(\tilde{\mathcal{A}}\tilde{\mathbf{a}}_1, \tilde{\mathcal{A}}\tilde{\mathbf{a}}_2)=\rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2)$, $\forall \tilde{\mathcal{A}} \in \widetilde{O(q)}$;

ii) $\rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) = \rho(\tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_1)$;

iii) $\rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) \geqslant 0$,而且$\rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) = 0 \Leftrightarrow \tilde{\mathbf{a}}_1 = \tilde{\mathbf{a}}_2$;

iv) 如果点$\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3$共线, 且$\Phi_0(\tilde{\mathbf{a}}_2)$位于点$\Phi_0(\tilde{\mathbf{a}}_1)$和$\Phi_0(\tilde{\mathbf{a}}_3)$之间, 那么$\rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_3) = \rho(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) + \rho(\tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3)$.

为了证明最后的这个断言, 需要验证等式

$$\Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_3) = \Delta(\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2) \cdot \Delta(\tilde{\mathbf{a}}_2, \tilde{\mathbf{a}}_3),$$

而这可以由重比的定义直接导出来.

值得注意的是这样一个事实, 罗巴切夫斯基空间不是有界的, 尽管它可以被“置于”球体之中.

罗巴切夫斯基度量允许在空间$\tilde{\Lambda}$中引进角度的概念. 以点$\tilde{\mathbf{e}}_0$ ($\dot{e}_0$是球面(6)的中心)为顶点的角同样也要在仿射图中计算. 为了定义以其他任意点$\tilde{\mathbf{e}}$为顶点的角$\mathcal{Y}$的量, 我们研究运动$\tilde{\mathcal{A}} \in \widetilde{O(q)}$, 它把$\tilde{\mathbf{e}}$变成$\tilde{\mathbf{e}}_0$, 而且设角$\mathcal{Y}$的量等于角$\tilde{\mathcal{A}}(\mathcal{Y})$的量(对于后面这个量, 在罗巴切夫斯基几何和仿射图中是重合的, 因为它的顶点是$\tilde{\mathbf{e}}_0$). 这个定义与运动$\tilde{\mathcal{A}}$的选择无关. 实际上, 如果有另外某一个运动$\tilde{\mathcal{B}}$使得$\tilde{\mathcal{B}}\tilde{\mathbf{e}} = \tilde{\mathbf{e}}_0$. 那么, $\tilde{\mathcal{C}} = \tilde{\mathcal{B}}\tilde{\mathcal{A}}^{-1}$也是个运动, 而且$\tilde{\mathcal{C}}\tilde{\mathbf{e}}_0 = \tilde{\mathbf{e}}_0$. 于是$\tilde{\mathcal{B}} = \tilde{\mathcal{C}}\tilde{\mathcal{A}}$, 同时, $\tilde{\mathcal{B}}(\mathcal{Y}) = \tilde{\mathcal{C}}(\tilde{\mathcal{A}}(\mathcal{Y}))$. 根据定理2, $\tilde{\mathcal{C}}$可以由仿射图($\mathbb{E}_0$, Φ_0)的欧几里得运动来实现. 所以, 角$\tilde{\mathcal{B}}(\mathcal{Y})$和角$\tilde{\mathcal{A}}(\mathcal{Y})$的量在仿射图中是一致的.

再次设$\mathcal{Y}$是个以$\tilde{\mathbf{e}}$为顶点的角, $\tilde{\mathcal{A}}$是任意一个罗巴切夫斯基运动, 且$\mathcal{Y}' = \tilde{\mathcal{A}}(\mathcal{Y})$. 按照定义, 角$\mathcal{Y}'$在$\tilde{\Lambda}$中的量等于角$\tilde{\mathcal{A}}'(\mathcal{Y}')$在仿射图中计算的量(欧几里得量), 其中$\tilde{\mathcal{A}}'$是把角$\mathcal{Y}'$的顶点变成$\tilde{\mathbf{e}}_0$的一个$\tilde{\Lambda}$运动. 但是, 这个时候, 运动$\tilde{\mathcal{A}}'\tilde{\mathcal{A}}$把$\tilde{\mathbf{e}}$变成$\tilde{\mathbf{e}}_0$, 从而角$\mathcal{Y}$的量等于在仿射图中计算的角$(\tilde{\mathcal{A}}'\tilde{\mathcal{A}})(\mathcal{Y})$的量. 因为

$$(\tilde{\mathcal{A}}'\tilde{\mathcal{A}})(\mathcal{Y}) = \tilde{\mathcal{A}}'(\tilde{\mathcal{A}}(\mathcal{Y})) = \tilde{\mathcal{A}}'(\mathcal{Y}'),$$

所以, 我们就证明了, 在罗巴切夫斯基运动之下角度是不变的.

在欧几里得空间里, 对于具有公共顶点的角成立的命题, 在罗巴切夫斯基空间中亦然成立. 特别地, **平角**等于π. 如果角$\dot{p}\dot{q}\dot{r}$和$\dot{r}\dot{q}\dot{s}$处于同一平面, 而且具有公共的边$\dot{q}\dot{r}$, 它在$\dot{q}\dot{p}$和$\dot{q}\dot{s}$之间, 那么, 它们的和等于角$\dot{p}\dot{q}\dot{s}$.

4. 罗巴切夫斯基平面 不难看出, 1维的罗巴切夫斯基几何与1维的欧几里得几何是一致的(如果$\tilde{\Lambda}$由不等式$x^2 < 1$给定, 那么, 函数

$$f(x) = \log \frac{1-x}{1+x}$$

决定了实直线上的映射$\tilde{\Lambda}$, 同时, $\tilde{\Lambda}$距离变成了欧几里得距离, $\tilde{\Lambda}$运动就是欧几里得运动).

我们可以停下来更方便地研究一下2维罗巴切夫斯基空间. 由第4章, 我们知道, 所有的欧几里得平面上以$\dot{e}$为不动点的运动或者是绕点$\dot{e}$旋转一个角度, 或者是一条经过点$\dot{e}$的直线的反射. 相应的术语可以照搬到罗巴切夫斯基平面上.

下列的命题同样地完全是正确的

i) 等腰三角形的两个底角相等, 顶角的平分线垂直于底边并且把它分成两半.

ii) 检验三角形全等的三个法则: 按一边及其两个邻角;按两边及其夹角;按三个边.

iii) 三角形的外角大于任意一个与其不相邻的内角.

iv) 在三角形中, 大边对大角.

v) 在三角形中, 任意一边小于另外两边之和.

vi) 由不在直线l上的一点$\tilde{\mathbf{e}}$可以引出一条到l的垂线, 而且是唯一的.

我们给出最后面的命题的论证. 为此, 研究$\tilde{\mathbf{e}}$和l在仿射图$(\mathbb{E}_\mathbf{e}, \Phi_\mathbf{e})$上的刻画. 为了简便起见, 有时会省略符号$\Phi_\mathbf{e}$.

直线l在这个图上可表示成圆Λ的弦, 而点$\tilde{\mathbf{e}}$表示成这个圆的中心$\dot{e}$, 如图25, 我们由$\dot{e}$在l上引出一欧几里得垂线h, 而且可以指出, 它在罗巴切夫斯基几何学的意义下也是垂直的. 为此, 研究平面$\tilde{\Lambda}$对于h的反射, 在图$\mathbb{E}_\mathbf{e}$上它怎样表示成欧几里得反射且因此把直线l变成自己. 由此可以得到, 在罗巴切夫斯基几何中, $h \perp l$.

垂线的唯一性, 如同在欧几里得几何中一样, 可以由命题引推出来. □

利用6), 不难相信, 垂线短于斜线, 作为习题, 我们设想去寻求点(a, b)到x轴的$\tilde{\Lambda}$距离(在圆$x^2 + y^2 < 1$之内)并写出与x轴等距点的几何轨迹的方程式.

这样一来, 很多2维欧几里得几何学中的定理对于2维的罗巴切夫斯基几何学也是对的. 但是, 在欧几里得平面上, 所有的关于图形构形的定理都可以由不多的几个几何公理推导出来, 这些公理表达了点和直线的最简单的性质(例如, 经过任意两点必可引出唯一的一条直线, 直线上任意三个点必有而且只有一个点位于另外两个点之间). 可以试着去验证, 这些公理中哪些在罗巴切夫斯基几何中仍保持正确. 如果能够证明, 它们所有的都是对的, 那么, 就可以断言, 欧几里得几何学中的所有定理

对于罗巴切夫斯基几何学也都是对的, 因为这些定理都是公理的逻辑性推论的结论. 而实际上, 罗巴切夫斯基几何学服从欧几里得几何学的所有公理, 只排除了一个, 就是平行公理, 它要求, 经过直线l外的一点$\dot{e}$最多只能引出一条与l不相交的直线. 但是, 由图26可以看出, 平行公理在罗巴切夫斯基几何学中不成立. 罗巴切夫斯基几何的存在性和无矛盾性是要证明的, 平行公理不能由欧几里得几何学中的原有的公理派生出来.

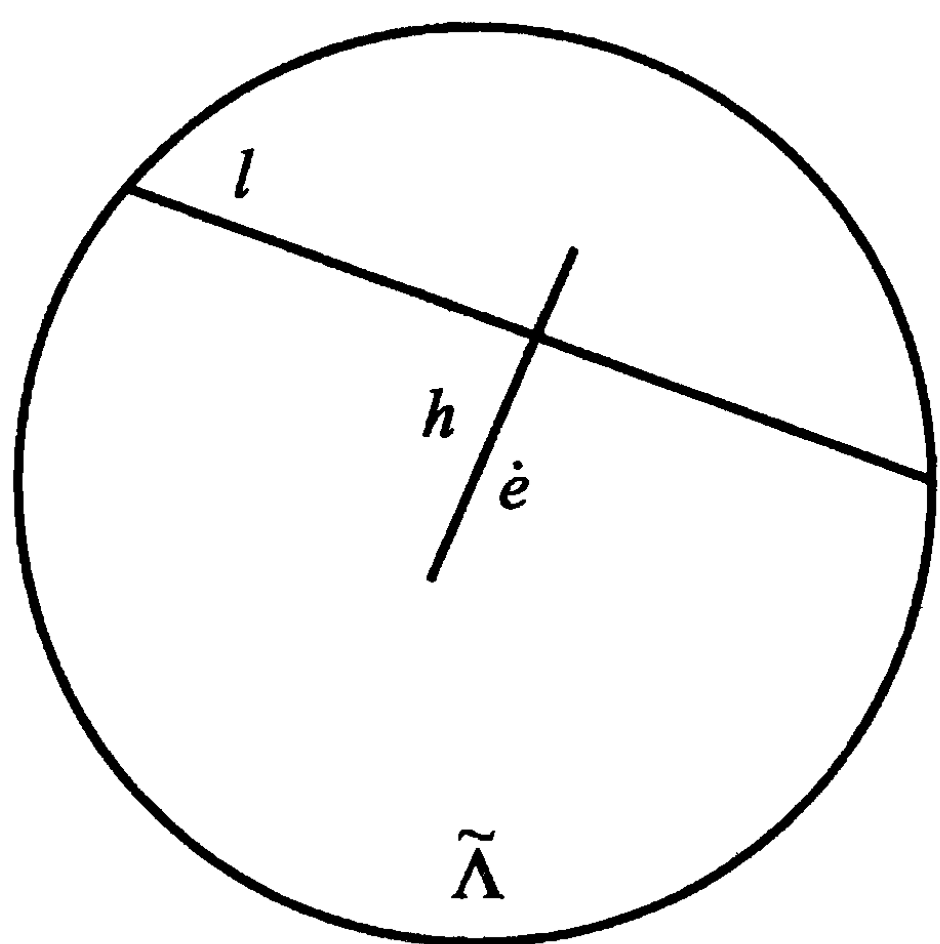


图25

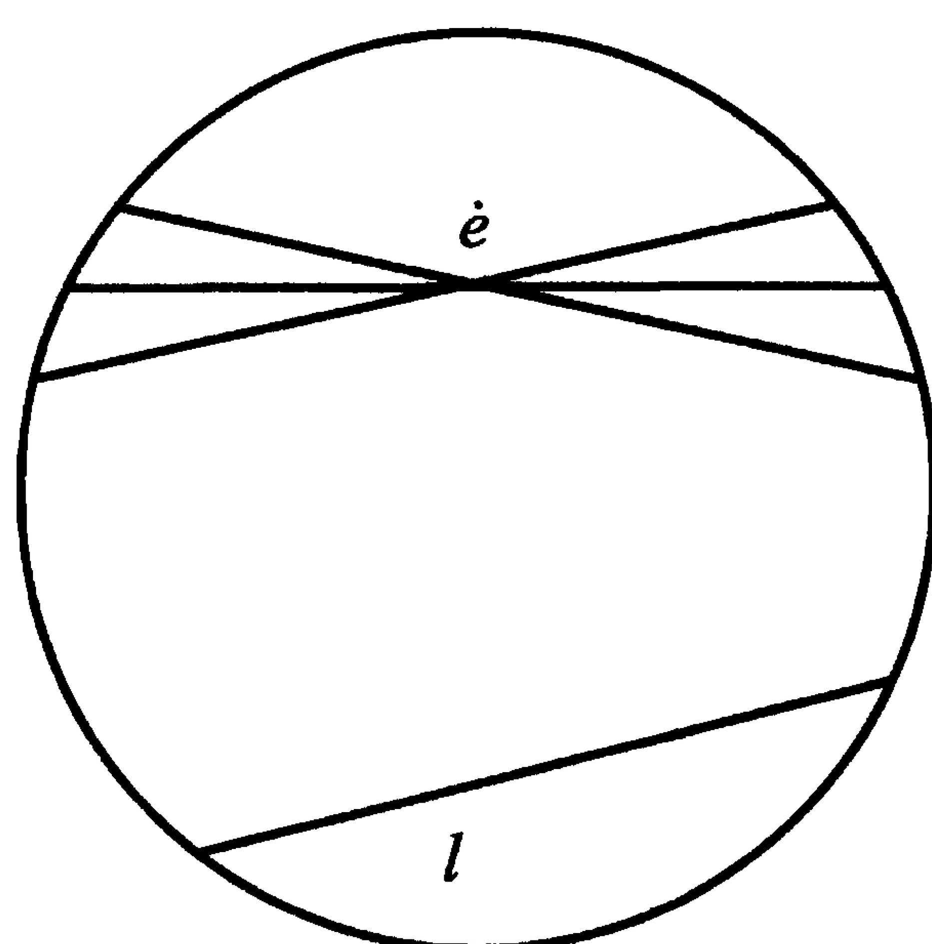


图26

罗巴切夫斯基几何学中具有本质上非欧几里得内涵的定理是

定理4　*罗巴切夫斯基平面上三角形的三个内角之和小于π.*

证明　i) 开始, 我们先来研究直角三角形, 它位于一个仿射图$(\mathbb{E}_0, \Phi_0)$上. 为简便起见(直到本节末), 在$\tilde{\Lambda}$中的点和在Λ中的点都用同一个大写的拉丁字母表示, 而把$\tilde{\Lambda}$运动用希腊字母表示. 借助于$\tilde{\Lambda}$运动, 我们可以使直角处顶点与点O重合. 这样就得到三角形AOB. 可以证明, 在罗巴切夫斯基几何意义下, 每个角$\widehat{BAO}$和$\widehat{OBA}$(记为$\widehat{BAO}|_\Lambda$)都比在仿射图中的度量(记为$\widehat{BAO}|_\mathbb{E}$)要小:

$$\widehat{BAO}|_\Lambda < \widehat{BAO}|_\mathbb{E}, \widehat{OBA}|_\Lambda < \widehat{OBA}|_\mathbb{E} \tag{10}$$

我们知道, $\widehat{AOB}|_\Lambda = \widehat{AOB}|_\mathbb{E} = \pi/2$. 由(9)就得到所需要的命题, 因为在欧几里得空间中三角形的三个内角之和等于π.

(9) 中的两个不等式的证明是一样的. 因而我们只限于证明第一个.

找一个运动把A变成O, 并用另外一个运动绕O点旋转, 使得边$\varphi(A)$. $\varphi(O)$ (φ是合成的运动)和边AO位于同一条直线上.

进一步的推论将在图27和这一目最初刻画的命题i)— vi)上进行.

按定义

$$\widehat{BAO}|_\Lambda = \widehat{\varphi(B)O\varphi(O)}|_\mathbb{E},$$

而且我们只要证明, 在仿射欧几里得图形内有不等式$\widehat{\varphi(B)O\varphi}(O) < \widehat{BAO}$就够了. 同样, 这个不等式因为有命题iv)就可以由欧几里得关系式

$$O\varphi(O) = AO, \tag{11}$$

$$\varphi(O)\varphi(B) < OB \tag{12}$$

得到.

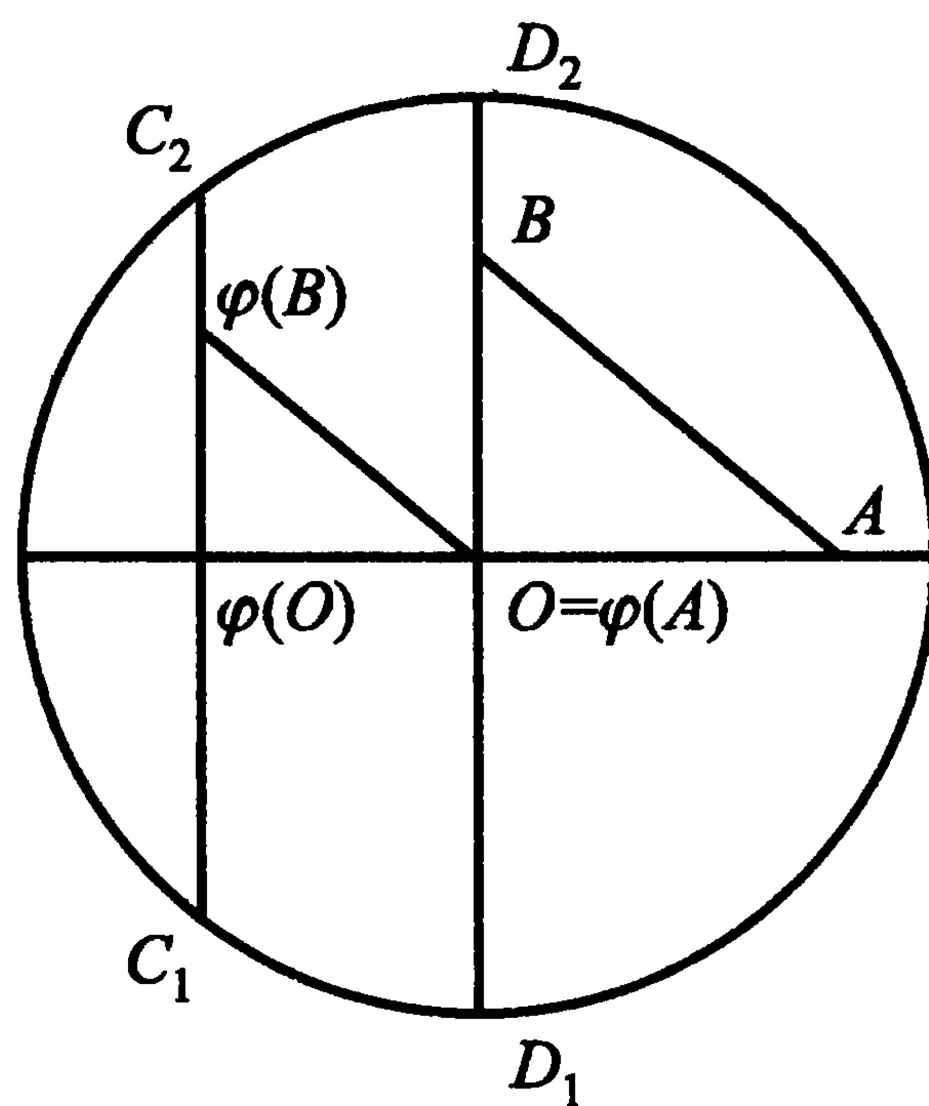


图27

线段AO和$O\varphi(O) = \varphi(A)\varphi(O)$在罗巴切夫斯基几何中是相等的, 因为一个可由另外一个经$\widetilde{\Lambda}$运动φ得到. 所以, 绕O旋转π角应该使A和$\varphi(O)$重合. 但是, 这个旋转就是一个欧几里得运动;于是就推出了等式(11).

为了证明等式(12), 利用罗巴切夫斯基距离的性质

$$\rho(O, B) = \rho(\varphi(O), \varphi(B)).$$

我们与图27相对应, 把欧几里得长度表示成$OB = x$, $\varphi(O)\varphi(B) = y$, $\varphi(O)C_2 = r$. 还要注意, $OD_i = 1$, $i = 1, 2$. 按照罗巴切夫斯基距离的定义

$$\begin{aligned}\rho(O, B) &= \log[O, B, D_2, D_1]\\ &= \log\left(\frac{OD_2}{D_2B} : \frac{OD_1}{D_1B}\right) = \log\left(\frac{1}{1-x}\cdot\frac{1+x}{1}\right) = \log\left(\frac{1+x}{1-x}\right),\end{aligned}$$

$$\begin{aligned}\rho(\varphi(O), \varphi(B)) &= \log[\varphi(O), \varphi(B), C_2, C_1]\\ &= \log\left(\frac{\varphi(O)C_2}{C_2\varphi(B)} : \frac{\varphi(O)C_1}{C_1\varphi(B)}\right) = \log\left(\frac{r}{r-y}\cdot\frac{r+y}{r}\right) = \log\left(\frac{r+y}{r-y}\right).\end{aligned}$$

由此可见,

$$\frac{r+y}{r-y} = \frac{1+x}{1-y} \Rightarrow \frac{1+y/r}{1-y/r} = \frac{1+x}{1-x} \Rightarrow y = rx,$$

又因为$r < 1$, 所以$y < x$.

ii) 我们利用命题iii) 来证明一般情形, 由它可以推出, 三角形不可能有两个钝角: 如果在$\triangle ABC$(图28)中在A, B两处的顶角都是钝角, 那么, 在顶角A处的外角就应该是锐角, 这意味着要违背命题iii), 因为它小于顶点B处的内角.

现在, 设在$\triangle ABC$中, 在顶点A和B处的内角都是锐角(如果其中之一是直角, 则定理已证(见前面的i)). 运用vi)由C向边AB引出一条垂直线CD(见图29). 那么, D必

然在A和B之间. 设若A位于D与B之间, 我们就会导出一个事实, $\triangle ABC$在顶点A的外角是个锐角, 而在顶点D的内角是个直角, 这就又与命题iii)相矛盾.

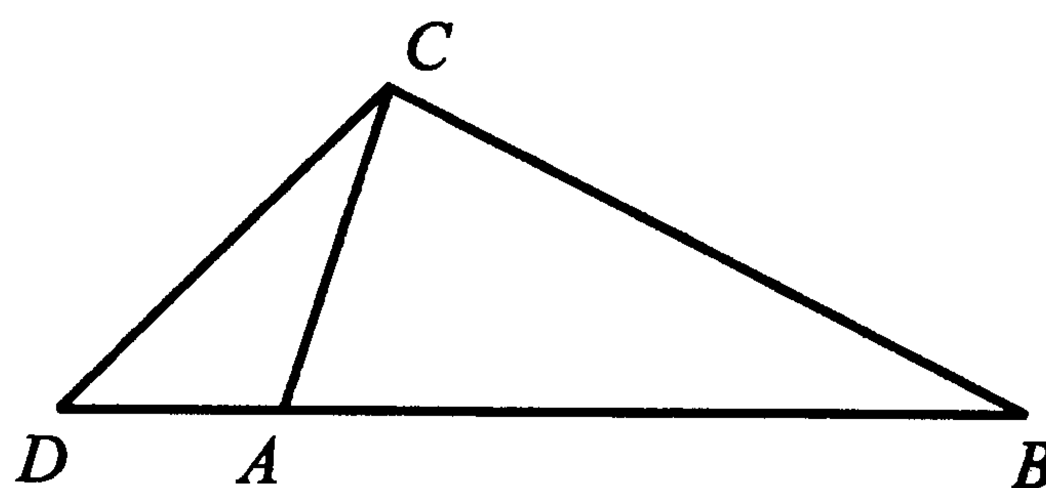


图28

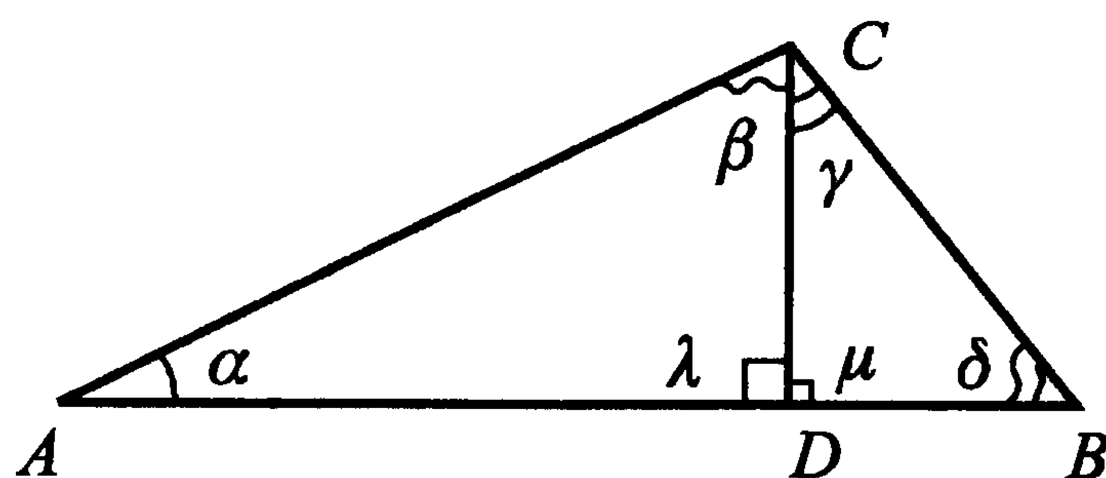


图29

这就是说, D位于A和B之间, 而三角形ABC被分成两个直角三角形ADC和BDC, 对它们每一个而言, 定理都是已经被证明了的. 我们有

$$\alpha + \beta + \lambda < \pi, \quad \gamma + \delta + \mu < \pi,$$

从而

$$\alpha + (\beta + \gamma) + \delta + \lambda + \mu < 2\pi,$$

同时, 又因为$\lambda = \mu = \pi/2$, 所以

$$\alpha + (\beta + \gamma) + \delta < \pi,$$

而这就是本定理的断言. □

说明 可以证明, 在罗巴切夫斯基平面上, 对任意小的$0 < \varepsilon < \pi$, 都存在一个三角形, 其内角之和为ε. 其次, 在欧几里得几何中, n边形的内角之和等于$\pi(n-2)$. 因为以α_1, α_2, $\cdots$, α_n 为内角的多边形M可以被分划成三角形, 所以, 由定理4, 在罗巴切夫斯基几何中, 不等式

$$\sum_{i=1}^{n} \alpha_i < \pi(n-2) \tag{13}$$

成立.

定理5 设

$$v(M) := \pi(n-2) - \sum_{i=1}^{n} \alpha_i.$$

如果

$$M = M_1 \cup M_2 \cup \cdots \cup M_k,$$

其中多边形M_1, M_2, $\cdots$, M_k双双均没有公共的内点, 那么,

$$v(M) = \sum_{i=1}^{k} v(M_i).$$

证明 显然，如果把多边形M某一个边上的某一个内点看成是外点，那么，边数就增加了一个，而内角就增加了2π，但$U(M)$没有变. 从而，可以认为，如果多边形M_i的顶点属于M_j，那么，就可以认为这个点是M_j的一个顶点. 于是，任意两个多边形M_i，M_j的交集就是连接它们的公共的边. 同样，不难看出，可以找到一个多边形M_j使得$\bigcup\limits_{i\neq j} M_i$，仍然是多边形. 这就允许我们把证明归结为$k=2$的情形.

设多边形M_1和M_2分别有n_1和n_2条边，其中有m条是公共的. 这些边是一个接一个的，否则$M_1 \cup M_2$就不是多边形了，多边形M的边数等于n_1+n_2-2m. 多边形M_1和M_2的内角之和等于多边形M的内角之和再减去$2\pi(m-1)$. 初等的计算就可表明

$$v(M) = v(M_1) + v(M_2). \qquad \square$$

函数v的性质类似于在欧几里得平面上的图形的面积的性质. 在罗巴切夫斯基平面中，值$v(M)$可以按定义看成是多边形M的面积.

我们来推导一个独特的关于判别三角形全等的准则.

定理6 *在罗巴切夫斯基平面上，如果三角形ABC的所有的角都对应地等于三角形$A'B'C'$的角，那么，这两个三角形全等.*

证明 罗巴切夫斯基平面上的运动可以把三角形$A'B'C'$的顶点A'与A重合，且进而使得边$A'B'$沿着AB走，而边$A'C'$沿AC走(图30). 边BC和$B'C'$不可能相交，否则就将与断言3)相矛盾.

于是，这两个三角形必然有一个整个地位于另一个三角形的内部(图31). 我们用M代表补余的四边形. 因为在$\triangle ABC$和$\triangle A'B'C'$中内角之和相等，所以，由定理5推出$v(M)=0$，但这与不等式(12)矛盾. 剩下的唯一的可能就是ABC和$A'B'C'$重合. $\square$

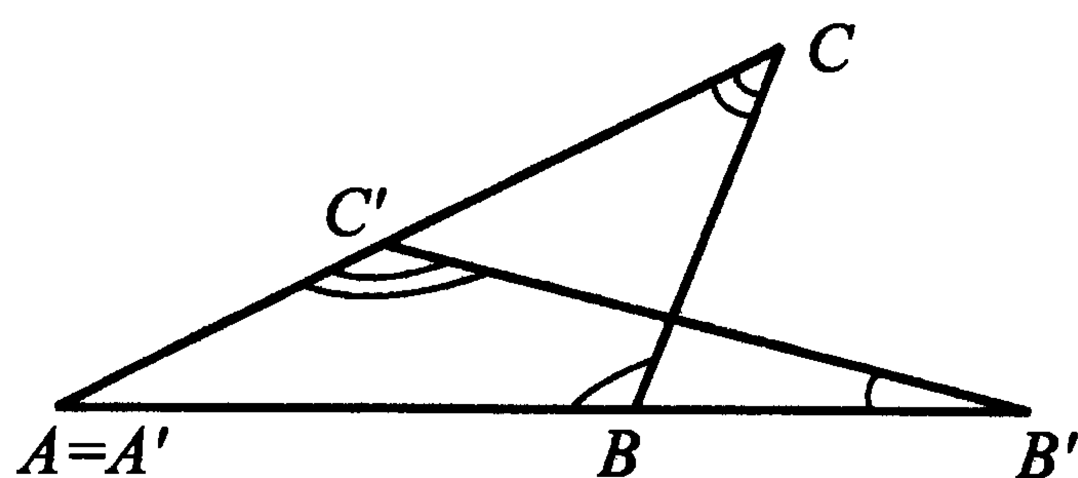


图30

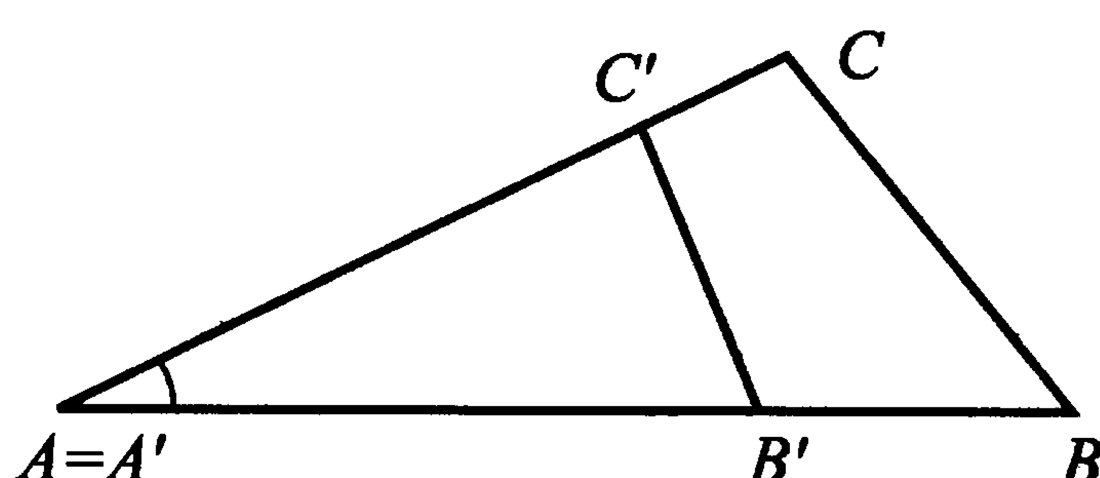


图31

有用的资料

[1] И. Р. 沙法列维奇的讲义记录稿. 20世纪60年代.

[2] Э. Б. 宾别尔哥. 线性代数与几何. 莫斯科大学出版, 1966.

[3] В. В. 尼库林, И. Р. 沙法列维奇. 群与几何. 科学出版社, 1983.

在书[3]的第Ⅳ章不仅对罗巴切夫斯基平面(在它的另外一种模型中)，而且对几何学的一般观点都会得到基础性的认知.

§6 有待解决的问题

1. 施特拉辛问题 正如在高斯量为Γ_n的情形(见[BA Ⅰ]第1章§3的注记2)一样, 对于大阶数n的两个方阵的乘积, 其必要的运算量的估计是有重要意义的. 因为数的乘法比加法要花费更多的力气, 所以在进行必要的机时先验估计的时候, 大多数的值与乘法运算的数量有关.

可以直接看出来, 两个$n\times n$阶矩阵的乘法需要做矩阵系数的n^3个乘法运算和$n^2(n-1)$个加法. 从$n=2$开始:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\cdot\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} aA+bB & aB+bD \\ cA+dC & cB+dD \end{pmatrix}.$$

由等式

$$\begin{aligned}
&aA+bC=(a-d)(A-D)+(b-d)(C+D)+d(A+C)+(a-b)D,\\
&aB+bD=-(a-b)D+a(B+D),\\
&cA+dC=(c-d)A+d(A+C),\\
&cB+dD=(a-d)(A-D)-(a-c)(A+B)+a(B+D)-(c-d)A
\end{aligned}$$

可以看出, 对于带有任意系数(它们本身也可能就是矩阵)的2×2阶矩阵的相乘, 只要依赖更多的18次加法(代替原有的4次), 则7次乘法就足够了(代替原有的8次). 把这个方法用到$n=2^k$阶的矩阵上, 并把它写成四个$2^{k-1}\times 2^{k-1}$的小块, 且简单地对k用归纳法, 就不难相信, 它们可以相乘, 且可用7^k次乘法和$6(7^k-4^k)$次加法完成. 现在, 设n是任意一个大自然数. 用数字0把n阶矩阵扩充成一个2^k阶的矩阵, 可以相信, 对于它们的乘法, 运算$O(n^{\log_2 7})=O(n^{2,81})$次就足够了(施特拉辛)(V. Strassen, 1969).

最近(Coppersmith-Winograd, 1982)证明了, 只要$O(n^{2,50})$次就够了. 剩下的是下面的一个未解决的问题:

猜想 存在这样的算法, 对于大自然数n, 它能保障$n\times n$阶矩阵的乘积经$O(n^{2+\varepsilon})$次运算实现, 其中ε是充分小的实数.

2. 正交分解 设$\mathcal{L}$是所有迹为0的n阶复矩阵的向量空间:

$$\mathcal{L}=\langle A\in M_n(\mathbb{C})\mid \mathrm{tr}A=0\rangle.$$

对于换位运算

$$[A,B]=AB-BA,$$

空间$\mathcal{L}$是个李代数$\mathfrak{sl}(n,\mathbb{C})$(或者, 简单地说, A_{n-1}型李代数), 在$\mathcal{L}$上定义非退化的纯量乘积

$$(A|B)=\mathrm{tr}\,AB,$$

它具有结合性(见第6章§2的(7)式):

$$([A,B]|C)=(A|[B,C]).$$

容易验证, 在所有的对角矩阵构成$\mathcal{L}$的子空间

$$\mathcal{H}_0=\langle \mathrm{diag}(\lambda_1,\cdots,\lambda_n)\mid\sum\lambda_i=0\rangle$$

上的限制$(A|B)|_{\mathcal{H}_0}$是非退化的. 对于它的任意一个共轭子代数(李代数理论中说是嘉当子代数$\mathcal{H}$).

$$\mathcal{H}=X^{-1}\mathcal{H}_0X,\quad \det X\neq 0$$

也是一样的, 可以证明, 作为向量空间, $\mathcal{L}$可以分解成嘉当子代数的直和

$$\mathcal{L}=\mathcal{H}_0\oplus\mathcal{H}_1\oplus\cdots\oplus\mathcal{H}_n, \tag{1}$$

其中

$$\mathcal{H}_i=X_i^{-1}\mathcal{H}_0X_i,$$

X_i是用适当的方法得到的所有的非退化矩阵. 现在, 提出如下的

问题 *能否找到那样的配极矩阵X_i, 使得分解式(1)中的子空间$\mathcal{H}_i$是两两正交的?*

如果能够如此, 那么, 就称这是李代数$\mathcal{L}$的一个**正交分解**(简记为OP). 从有限群分解的整格等等的观点来看, OP的存在性是很吸引人的(见关于这方面的书籍: A. I. Kostrikin (即本书的作者), Pham Huu Tiep. Orthogonal Decompositious and Integral Lattices, BerLin-New York: Walter de Gruyter, 1994).

猜想 *李代数$\mathcal{L}=\mathfrak{sl}(n,\mathbb{C})$的OP存在的充要条件是$n$为某个素数$p$的方幂, 即$n=p^k$.*

猜想的一个方面已经被证明了: 对任意的$n=p^k$, OP是可以建立起来的, 剩下的是要证明, $n\neq p^k$就不能建立OP. 显然, $n=6$是最小的这样的数. 对它, 猜想或者可证明或者被推翻. 换言之, 在$\mathfrak{sl}(6,\mathbb{C})$中存在7个两两正交的5维子空间$\mathcal{H}_i$满足条件(1)吗?要回答这个看起来好像十分具体的问题, 迄今为止, 尚未可能实现.

3. 有限射影平面 将第5章的§3加以扩展, 我们称点和可分离的子集, 即满足下列公理的直线的集合Π是一个**射影平面**:

P1 *任意两个不同的点必然属于而且只能属于一条直线;*

P2 *任意两条不同的直线必然包含而且只能包含一个公共点;*

P3 *存在四个点, 它们之中的任意三个点都不在同一条直线上.*

下列定理成立

定理1 *设整数$n\geqslant 2$已经给定. 在投影平面Π上, 下列性质等价:*

i) *有某条直线刚好由$n+1$个点组成;*

ii) 有某个点刚好属于$n+1$条不同的直线;

iii) 每一条直线刚好都由$n+1$个点组成;

iv) 每个点都刚好属于$n+1$个不同的直线;

v) Π包含n^2+n+1个点;

vi) Π包含n^2+n+1条直线.

证明, 比如, 可以从下面这本书摘取: The Theory of Group, Marshall. Hall. Jr. New York. The Macmillan Company, 1959 (俄译本: Теория групп, *Холл. М.*, М.: ИЛ, 1962).

定义 在$|\Pi|=n^2+n+1$的情形, 称**射影平面是n阶的**.

在[BA Ⅲ]中将判明, 对任意的素数p和任意的自然数k都存在一个域$\mathbb{F}_q=GF(q)$, 它含有$q=p^k$个元素(当$k=1$时, 我们已知). 如果V是$\mathbb{F}_q$上的3维向量空间, 那么$\mathbb{F}_q\mathbb{P}^2=\mathbb{P}(V)$是个$q$阶的射影平面的德萨格空间(见[2]). 一般说来, 当$k>1$时, q阶的非德萨格射影平面也是存在的. 但是(这本身也是很著名的), 到目前为止, 还未能建立一个阶数$n\neq p^k$的射影平面.

猜想 任意有限的射影平面的阶数都应当是某个素数p的方幂.

按着这个猜想本身的叙述容易令人想起第2目的那个猜想, 以及实际上存在怎样的暗藏着的类似结果. 但是, 射影平面问题提出来更为久远, 然而成果却是相当地微不足道.

我们引进一个有趣的算术特征的结果(Bruck, Ryser): 如果Π是个n阶射影平面, 而$n\equiv 1,2(\mathrm{mod}\ 4)$, 那么, $n=a^2+b^2$是两个整数的平方和.

比方说, $n=6$就不能被表达成两个整数的平方和, 所以, 就不可能是某个射影平面的阶数. 长期以来, 关于建立$10=3^2+1^2$阶的射影平面的问题, 变成了一个有待解决的问题. 仅在电子计算机" Крейг"上粗略估计, 经过整整700小时过程才得到了结论, 不存在那样的射影平面.

沿同样的道路走下去还是没有可能. 对于$n>10$的猜想依然没能得到证明.

4. 空间的基底与拉丁方 在理解了第一章的水平上可以叙述如下的:

猜想1(Rota, 1989) 设V是任意一个无限域k上的一个n维向量空间. 再设$B_1, B_2,\cdots,B_n$是空间V的某一类共计n个基底.

那么, 可以把每一个基底中的向量调整顺序, 比方说, 已经变成$B_i=(\mathbf{b}_{i1},\mathbf{b}_{i2},\cdots,\mathbf{b}_{in})$, 使得向量组$C_j=(\mathbf{b}_{1j},\ \mathbf{b}_{2j},\ \cdots,\ \mathbf{b}_{nj})$, $1\leqslant j\leqslant n$, 也都是V的基底.

换言之, 经过适当的调整以后, 使得$B=(\mathbf{b}_{ij})$的每一横行的元素构成一个基底, 每一个纵列的元素也构成一个基底. 当$n=3$时, 这是小孩子们的一个练习题. 而对于任意n的猜想. 它与不变性理论有关, 现在仍然未被证明. 在这中间, 已经确立了, 对于任意偶数的情形, 猜想1可以由下面的猜想2推导出来, 而猜想2与另一个古老的组合课题有关系.

称用n个符号, 比方说, 整数0, 1, $\cdots$, $n-1$, 填满的$n\times n$平方表L(或者矩阵)是

一个**拉丁方**, 如果, 在它的每一行上的元素都是不同的, 同时每列上的各元素也是不同的. 对应于给定一行或者给定一列, 对应的集合$\{0,1,\cdots,n-1\}$上的这个置换的符号就说是拉丁方L的这一行或者这一列的符号. 所有的这$2n$个行的符号和列的符号的乘积被称为是**拉丁方L的符号**, 记为$\varepsilon(L)$. 按定义, 当$\varepsilon(L)=+1$时, 称**L是偶的**, 当$\varepsilon(L)=-1$时, 称L是**奇的**. 当n为奇数时, n阶的偶拉丁方的个数与n阶奇拉丁方的个数是一样的. 但, 即使是$n=2$, 4, 6, 这个事实都不对.

猜想2 (Alon-Tarsi, 1986)　*设n是个偶自然数. 那么, $\sum\varepsilon(L)\neq 0$, 其中的求和取遍所有的n阶拉丁方.*

前不久, 彻底证明了, 如果$n=p+1$, 其中p是个奇素数, 那么, 猜想2是正确的[Drisko A. A//Advances in Math. —1997. —No. 128—P. 20—35], 从而对任意$n=p+1$, 猜想1是正确的, 可以把这个结果推广到一般的$n=p^k+1$的情形吗?

习题解答与提示

数码p, q, r是指第p章§q的习题r.

1.2.9 $\dim \mathrm{Mag}_3(\mathbb{Q})=3$, $\dim \mathrm{Mag}_4(\mathbb{Q}) = 8$.

1.2.10 直接验证, 对任意的半幻矩阵A都有$SA = \sigma(A)S = AS$. 此外, 对任意指数$m \geqslant 1$ 都有$S^m = n^{m-1}S$.

现在, 设$\mathrm{Mag}_n^0(\mathbb{Q})$(相应地$\mathrm{SMag}_n^0(\mathbb{Q})$)是所有的迹为零的幻方矩阵(相应地半幻矩阵)的集合, 那么

$$\mathrm{Mag}_n(\mathbb{Q}) = \mathrm{Mag}_n^0(\mathbb{Q}) \oplus \mathbb{Q}S, \mathrm{SMag}_n(\mathbb{Q}) = \mathrm{SMag}_n^0(\mathbb{Q}) \oplus \mathbb{Q}S \tag{$*$}$$

(只要注意到这样事实就够了, 即$\mathrm{tr}\left(A - \dfrac{1}{n}\mathrm{tr}(A)S\right) = 0$).

其次,

$$\dim \mathrm{Mag}_n^0(\mathbb{Q}) \geqslant \dim \mathrm{SMag}_n^0(\mathbb{Q}) - 2,$$

因为幻方的空间是由半幻方再附加两个限制后得到的, 更准确些,

$$\mathrm{SMag}_n^0(\mathbb{Q}) = \mathrm{Mag}_n^0(\mathbb{Q}) \oplus \mathbb{Q}E \oplus \mathbb{Q}D. \tag{$**$}$$

实际上, 我们设, 有关系式$A + \lambda E + \mu D = 0$, 其中λ, $\mu \in \mathbb{Q}$, $A \in \mathrm{Mag}_n^0(\mathbb{Q})$. 用S乘它就得到$(\lambda + \mu)S = 0$, 从而$\lambda + \mu = 0$. 而同样地, 由迹, 我们又得到$n\lambda + \mu\mathrm{tr}\, D = 0$. 必有结论$\lambda = \mu = 0$.

剩下来只需将等式($*$)和($**$)加以对照.

1.2.11 提示 考察由子空间V_i 的直和到空间V 的一个$k-1$分量的直和上的映射, 即

$$(\mathbf{v}_1, \cdots, \mathbf{v}_k) \mapsto (\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_2 - \mathbf{v}_3, \cdots, \mathbf{v}_{k-1} - \mathbf{v}_k).$$

所有的V_i 的交集是这个映射的核, 它的维数不小于$\dim V_1 + \cdots + \dim V_k - (k-1)n$.

1.4.3 $-1/2 < \lambda < 1, \mu < -2$.

1.4.4 $z_1 = x_1y_1 + x_2y_3 + x_3y_2$, $z_2 = x_1y_2 + x_2y_1 + x_3y_3$, $x_1y_3 + x_2y_2 + x_3y_1$.

2.1.3 显然.

2.2.7 设$\dim \operatorname{Im} \mathcal{A}^{i-1} = k + l$, 且

$$\operatorname{Im} \mathcal{A}^{i-1} = \left\langle \mathcal{A}^{i-1}\mathbf{e}_1, \cdots, \mathcal{A}^{i-1}\mathbf{e}_k;\ \mathcal{A}^{i-1}\mathbf{e}_{k+1}, \cdots, \mathcal{A}^{i-1}\mathbf{e}_{k+l}\right\rangle,$$

其中$\langle\ \mathcal{A}^{i-1}\mathbf{e}_1, \cdots, \mathcal{A}^{i-1}\mathbf{e}_k\ \rangle = \operatorname{Im} \mathcal{A}^{i-1} \cap \operatorname{Ker} \mathcal{A}$, 这样一来,

$$V = \langle \mathbf{e}_1 \cdots, \mathbf{e}_k; e_{k+1} \cdots, \mathbf{e}_{k+l}; \mathbf{e}_{k+l+1}, \cdots, \mathbf{e}_n\rangle,$$
$$\langle \mathbf{e}_{k+l+1}, \cdots, \mathbf{e}_n\rangle = \operatorname{Ker} \mathcal{A}^{i-1},$$
$$\langle \mathbf{e}_1, \cdots, \mathbf{e}_k; \mathbf{e}_{k+l+1}, \cdots, \mathbf{e}_n\rangle = \operatorname{Ker} \mathcal{A}^i.$$

我们有

$$\dim\ \operatorname{Ker} \mathcal{A}^i = n - l, \quad \dim\ \operatorname{Ker} \mathcal{A}^{i-1} = n - k - l,$$
$$\dim\ (\operatorname{Im} \mathcal{A}^{i-1} \cap \operatorname{Ker} \mathcal{A}) = k = (n - l) - (n - k - l).$$

2.2.8 据条件, $A = C^{-1}BC$, 其中$C \in M_n(\mathbb{C})$. 需要证明, 在方程式$XA = BX$ 的解$X = C$ 中间能找到一个实矩阵解D. 我们设这个方程式的解在复数域$\mathbb{C}$ 上构成一个向量空间W , 它的基底是$C_1, \cdots, C_m$. 把C_j 表达成$C_j = G_j + iH_j$ 的形式, $G_j, H_j \in M_n(\mathbb{R})$. 可以相信, $G_jA = BG_j$且$H_jA = BH_j$, 也就是说, 允许向量空间W 有实的基底($D_1, \cdots, D_m$)(它可由$G_1, \cdots, G_m$; $H_1, \cdots, H_m$ 中选择出来). 设$f = (t_1, \cdots, t_m) = \det(t_1D_1 + \cdots + t_mD_m)$ 是m个变元的实多项式, 按条件, 它在域$\mathbb{C}$ 上不恒等于零, 于是它在实数域$\mathbb{R}$ 上也不恒等于零. 这意味着, 这个矩阵方程有非退化的实数解$t_1^0D_1 + \cdots + t_m^0D_m$.

2.2.9 (1) 是显然的.

(2) 当$n = \dim_{\mathfrak{K}} V = 1$ 时, 命题是真的. 我们对n用归纳法, 设W 是V 的任意一个超平面, 那么$V = \langle W, \mathbf{e}\rangle_{\mathfrak{K}}$. 按归纳法假定, 存在一个向量$\mathbf{w} \in W$ 使得
$\mu_{\mathcal{A},\mathbf{w}}(\mathcal{A})W = \mathbf{0}$, 我们设

$$f_k(t) = \mu_{\mathcal{A},\mathbf{w}+k\mathbf{e}}(t), \quad k = 0, 1, \cdots$$

依据1), $f_k(t)$ 应整除$\mu_{\mathcal{A}}(t)$. 但是, 这种因式只有有限个, 所以必有一对指标i, j, $i \neq j$使得$f_i(t) = f_j(t) := f(t)$ (域$\mathfrak{K}$是无穷的). 即

$$f(\mathcal{A})(\mathbf{w} + i\mathbf{e}) = \mathbf{0} = f(\mathcal{A})(\mathbf{w} + j\mathbf{e}).$$

从而$f(\mathcal{A})\mathbf{w} = \mathbf{0} = f(\mathcal{A})\mathbf{e}$, 同时$\mathbf{a} = \mathbf{w} + i\mathbf{e}$ 或者$\mathbf{a} = \mathbf{w} + j\mathbf{e}$ 即为所求.

2.2.10 (1) 显然,

$$(*) \Rightarrow W + U \supset W_1 + (U \cap V_1) = V_1.$$

进一步, $\mathcal{P}_2(W+U)=W_2+\mathcal{P}_2(U)=V_2$. 所以, $V=W+U$.

(2) 按条件, $\mathbf{v}\in V \Rightarrow \mathbf{v}=\mathbf{w}+\mathbf{u},\ \mathbf{w}\in W,\ \mathbf{u}\in U$. 这意味着, $\mathbf{v}=[\mathcal{P}_1(W)+\mathcal{P}_2(W)]+[\mathcal{P}_1(\mathbf{u})+\mathcal{P}_2(\mathbf{u})]=[\mathcal{P}_1(\mathbf{w})+\mathcal{P}_1(\mathbf{v})]+[\mathcal{P}_2(\mathbf{w})+\mathcal{P}_2(\mathbf{u})]$. 如果$\mathbf{v}\in V_1$, 那么$\mathcal{P}_2(\mathbf{w})+\mathcal{P}_2(\mathbf{u})=\mathbf{0}$. 但是, 据条件$\mathcal{P}_2(U)\wedge W_1=\mathbf{0}$. 所以, $\mathcal{P}_2(\mathbf{w})=\mathbf{0}=\mathcal{P}_2(\mathbf{u})$, 从而, $V_1=W_1+(U\wedge V_1)$. 把$\mathcal{P}_2$ 作用到$V=W+U$ 上, 我们得到$V_2=W_2+\mathcal{P}_2(U)$, 也就是, 分解式(*)成立.

最后, $\mathbf{w}\in W\cap U \Rightarrow \mathcal{P}_2(\mathbf{w})\in\mathcal{P}_2(W)\cap\mathcal{P}_2(U)=W_2\cap\mathcal{P}_2(U)=\mathbf{0}\Rightarrow \mathbf{w}\in W_1$ 且$\mathbf{w}\in W_1\cap U$, 也就是$W\cap U=W_1\cap U$.

2.2.11 线性算子$\mathcal{A}:V\to V$, $\mathrm{tr}\,\mathcal{A}=0$,对应矩阵$A$. 同时, 设$V=\langle\mathbf{e}_1,\cdots,\mathbf{e}_n\rangle$. 问题归结为, 构造某一个基底$(\mathbf{e}_1',\cdots,\mathbf{e}_n')$, 在这个基底之下, 算子$\mathcal{A}$ 对应的矩阵$\mathcal{A}'$ 满足所需的要求.

如果$\mathcal{A}=\lambda\mathcal{E}$, 那么, $\mathrm{tr}\,\mathcal{A}=n\lambda$ 且$\lambda=0$. 如果$\mathcal{A}\neq\lambda\mathcal{E}$, 那么, 必然可以找到两个线性无关的向量$\mathbf{e}_1',\mathbf{e}_2'=\mathcal{A}\mathbf{e}_1'$. 这就给出了算子$\mathcal{A}$ 的矩阵的左上角为零的可能性. 设, 已经构造出线性无关的向量$\mathbf{e}_1',\cdots,\mathbf{e}_k'$, 相对于它们, $a_{11}'=\cdots=a_{kk}'=0$. 现在, 我们取$\mathbf{e}_{k+1}'\notin V_k$, $V_k=\langle\mathbf{e}_1',\cdots,\mathbf{e}_k'\rangle$, 使得$\mathbf{e}_{k+2}':=\mathcal{A}\mathbf{e}_{k+1}'$ 不属于$\langle V_k,\mathbf{e}_{k+1}'\rangle$. 这样, 我们就在对角线上又得到一个零. 如果这样做是不可能的, 那就会出现这样的情形, 只要$\mathbf{x}\notin V_k$ 就必有$\mathcal{A}\mathbf{x}\in V_k$. 如果对任意$\mathbf{x}\notin V_k$ 都有$\mathcal{A}\mathbf{x}\in V_k$, 那么, 由$\mathbf{e}_1',\cdots,\mathbf{e}_k'$ 扩展成的空间V 的任意一个基底都即为所要找到的基底. 现在, 设$\mathcal{A}\mathbf{x}=\lambda_x\mathbf{x}+\cdots$, 其中省略处是一个V_k 中的向量. 如果对任意$\mathbf{x}\notin V_k$, $\lambda_{\mathbf{x}}=\lambda$, 那么, 又有$\mathrm{tr}\,\mathcal{A}=(n-k)\lambda=0\Rightarrow\lambda=0$, 就证完了. 如果对某一对$\mathbf{u},\ \mathbf{v}\notin V_k$有$\mathcal{A}\mathbf{u}=\lambda\mathbf{u},\ \ \mathcal{A}\mathbf{v}=\mu\mathbf{v},\ \ \lambda\neq\mu$, 那么, 我们设, 比方说$\mathbf{e}_{k+1}'=\mathbf{u}+\mathbf{v},\ \ \mathbf{e}_{k+2}'=\mathcal{A}(\mathbf{u}+\mathbf{v})=\lambda\mathbf{u}+\mu\mathbf{v}$. 这就给出在对角线上再增加一个零的可能性. 显然, 可以用归纳法完成推理.

2.2.12 是的, 存在. 如果, 比方说$n=p$ 且$A=E$, 那么$\mathrm{tr}\,A=0$, 而所有的相似矩阵$C^{-1}AC$ 都与A重合.

2.3.2 对矩阵E_{ij}, $1\leqslant i,j\leqslant n$, 我们设$P_i=\mathcal{D}(E_{ii})$. 于是

$$P_iP_j=\mathcal{D}(E_{ii})\mathcal{D}(E_{jj})=\mathcal{D}(E_{ii}E_{jj})=\mathcal{D}(\delta_{ij}E_{ii})=\delta_{ij}P_i,$$

也就是说, $P_1,\cdots,P_n$ 是等方矩阵的正交系. 在向量空间V 的一个基底之下把$n\times n$ 矩阵与算子等同起来, 对于单位矩阵E, 我们有

$$\mathcal{D}(E)V=P_1V\oplus\cdots\oplus P_nV,$$

从而, $\sum\limits_i\mathrm{rank}\,P_i=\mathrm{rank}\,\mathcal{D}(E)$. 如果对所有的i, 都有$P_i\neq0$, 那么, 可由此推出, $\mathrm{rank}\,P_i=1$. 我们当然期待是这样的. 实际上,在相反的情形, 对某个i有$P_i=0$, 那么,

$$P_j=\mathcal{D}(E_{ji}E_{ii}E_{ij})=\mathcal{D}(E_{ji})P_if(E_{ij})=0,\quad\forall j,$$

所以, $\mathcal{D}(E)=\sum\limits_iP_i=0$. 但是在这个情形$\mathcal{D}(X)=\mathcal{D}(EX)=\mathcal{D}(E)f(X)=0$ 对任意

一个矩阵X都成立, 这是一个矛盾, 我们就得出结论, $\mathcal{D}(E)=n$. 再由$\mathcal{D}(E)^2=\mathcal{D}(E)$. 得出, $\mathcal{D}(E)=E$.

其次, $\dim\ P_1V=1\Rightarrow P_1V=\langle A^{(1)}\rangle$. 设$A^{(i)}$ 是$V=\mathfrak{K}^n$中形如$A^{(i)}=\mathcal{D}(E_{i1})A^{(1)}$, $2\leqslant i\leqslant n$ 的列向量. 因为$E_{ij}E_{kl}=\delta_{jk}E_{il}$, 所以,

$$\mathcal{D}(E_{ij})A^{(k)}=\mathcal{D}(E_{ij})\mathcal{D}(E_{k1})A^{(1)}=\mathcal{D}(E_{ij}E_{k1})=\delta_{jk}A^{(i)}. \tag{$*$}$$

特别地, $\mathcal{D}(E_{1i})A^{(i)}=A^{(1)}$ 且$P_iA^{(i)}=A^{(i)}$. 这样一来, $A^{(i)}\neq 0$ 且$A^{(i)}\in P_iV$, 也就是, $P_iV=\langle A^{(i)}\rangle,\quad 1\leqslant i\leqslant n;\quad V=\langle A^{(1)},\cdots,A^{(n)}\rangle$. 设$A=(A^{(1)},\cdots,A^{(n)})$ 是$n\times n$阶矩阵(列向量排成行). 按定义, $\operatorname{rank} A=n$. 按照$(*)$

$$\mathcal{D}(E_{ij})A=(\mathbf{0},\cdots,\underbrace{A^{(i)}}_{j},\cdots,\mathbf{0})=AE_{ij},$$

而且, 这意味着, $\mathcal{D}(X)A=AX$. 令$C=A^{-1}$. 我们有$\mathcal{D}(X)=C^{-1}XC=f_C(X)$.

□

2.3.3 按定义

$$\mathcal{A}^pV=\mathcal{A}^{p+1}V=\cdots=\mathcal{A}^{2p}V=\mathcal{A}^p(\mathcal{A}^pV),$$

从而, $\operatorname{Im}\mathcal{A}^p\cap\operatorname{Ker}\mathcal{A}^p=\mathbf{0}$, 再加上§1的定理4, 就得到直和$V=\operatorname{Ker}\mathcal{A}^p\oplus\operatorname{Im}\mathcal{A}^p$. 两个加项对于$\mathcal{A}$ 的不变性是显然的且从前已经指出过了.

2.3.5 首一多项式$f(t)=t^m+\sum\limits_{i=1}^{m}a_it^{m-i}\in\mathbb{C}[t]$ 对应一个复共轭多项式$\overline{f}(t)=t^m+\sum\limits_{i=1}^{m}\bar{a}_it^{m-i}$. 由条件推知, $\chi_A(t)=f(t)\overline{f}(t)$ 是两个首一的互素的复共轭的多项式的乘积. 从而有互素的实多项式$p(t)$, $q(t)$ 使得$\chi_A=p^2+q^2$. 我们就能找到$r(t)$, $s(t)\in\mathbb{R}[t]$ 使得$pr-qs=1$, 我们再设$g=ps+qr$, $h=r^2+s^2$, 则

$$1+g^2=(pr-qs)^2+(ps+qr)^2=(r^2+s^2)(p^2+q^2)=h(t)\chi_A(t).$$

剩下来就是规定$B=g(A)$.

2.3.6 如果矩阵A,B中有一个是非退化的, 那么, 一切都对了. 比方说, $\chi_{AB}(\lambda)=\det(AB-\lambda E)=\det(A^{-1}(AB-\lambda E)A)=\det(BA-\lambda E)=\chi_{BA}$. 因为$\lambda$多项式对矩阵$A$, B的系数是连续的, 所以, 上面的等式对退化矩阵也成立.

2.3.7 把矩阵B化成对角形式.

2.3.8 可以把矩阵$A\in M_n(\mathbb{Q})$ 解释成为坐标空间$\mathbb{Q}^n$ 上的一个线性算子, 它是半幻的, 当且仅当, 列向量$\mathbf{e}:=[1,\ 1,\ \cdots,\ 1]$ 是A和${}^tA$ 的一个特征向量(对于同一个特征值$\lambda=\sigma(A)=\sigma({}^tA)$). 这就表明, $A,\ B\in\operatorname{SMag}_n(\mathbb{Q})\Rightarrow(AB)\mathbf{e}=A(B\mathbf{e})=\sigma(B)A\mathbf{e}=\sigma(A)\sigma(B)\mathbf{e}$; ${}^t(AB)\mathbf{e}={}^tB({}^tA\mathbf{e})=\sigma(A)\sigma(B)\mathbf{e}\Rightarrow AB\in\operatorname{SMag}_n(\mathbb{Q})$.

2.3.9 当$n=1$ 时命题为真. 然后, 对n用归纳法. 设n是使得命题不真的最小阶数, 也就是, 对于某个$n\times n$ 阶矩阵$A=(a_{ij})$, $S(A)$ 中的所有矩阵都具有一个特征

值1. 换句话说, 任意矩阵$D = \operatorname{diag}(\theta_1, \cdots, \theta_n)$, $\theta_i = \pm 1$, 都对应一个列向量$\mathbf{x} \neq \mathbf{0}$, 且$DA\mathbf{x} = \mathbf{x}$, 或者, 也就是$(A - D)\mathbf{x} = \mathbf{0}$. 我们将其化成关系式

$$d(\theta_1, \cdots, \theta_n) = \det(A - \operatorname{diag}(\theta_1, \cdots, \theta_n)) = 0$$

对$\theta_i, i = 1, \cdots, n$ 的2^n 种选择都成立. 按行列式的第1列展开, 我们得到表达式

$$d(\theta_1, \cdots, \theta_n) = (a_{11} - \theta_1)d^*(\theta_2, \cdots, \theta_n) + \text{不含}\theta_1\text{的项}. \tag{$*$}$$

按归纳法假设, $n-1$ 阶行列式$d^*(\theta_2, \cdots, \theta_n)$ 必可选出在$\theta_2^0, \cdots, \theta_n^0$ 处不为零. 这时, 由

$$d(1, \theta_2^0, \cdots, \theta_n^0) = 0 = d(-1, \theta_2^0, \cdots, \theta_n^0)$$

和$(*)$即可推出

$$(a_{11} - 1)d^*(\theta_2^0, \cdots, \theta_n^0) = (a_{11} + 1)d^*(\theta_2^0, \cdots, \theta_n^0),$$

从而$a_{11} - 1 = a_{11} + 1$, 这是一个矛盾, 因为char $\mathfrak{K} \neq 2$.

2.3.10 蕴含式的一个方向可由第1目中关于投影的结果得到. 现在, 我们设, rank $\mathcal{A} = r$, rank $(\mathcal{E} - \mathcal{A}) = n - r$. 于是$V = \langle \mathbf{e}_1, \cdots, \mathbf{e}_r; \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle$, 向量$\mathcal{A}\mathbf{e}_1, \cdots, \mathcal{A}\mathbf{e}_r$ 线性无关, 且$\mathcal{A}\mathbf{e}_{r+1} = \cdots = \mathcal{A}\mathbf{e}_n = \mathbf{0}$. 因为$(\mathcal{E} - \mathcal{A}) \langle \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle = \langle \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle$ 且rank$(\mathcal{E} - \mathcal{A}) = n - r$, 所以

$$(\mathcal{E} - \mathcal{A}) \langle \mathbf{e}_1, \cdots, \mathbf{e}_r \rangle \subseteq \langle \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle.$$

如果$(\mathcal{E} - \mathcal{A})\mathbf{e}_i = \alpha_{r+1}\mathbf{e}_{r+1} + \cdots + \alpha_n\mathbf{e}_n$, 那么, 只要令

$$\mathbf{e}_i' = \mathbf{e}_i - \alpha_{r+1}\mathbf{e}_{r+1} - \cdots - \alpha_n\mathbf{e}_n, \quad i \leqslant r; \quad \mathbf{e}_i' = \mathbf{e}_i, \quad r+1 \leqslant i \leqslant n,$$

我们就有$\mathcal{A}\mathbf{e}_i' = \mathcal{A}\mathbf{e}_i$ 而且

$$(\mathcal{E} - \mathcal{A})\mathbf{e}_i' = \mathbf{0}, \quad i \leqslant r; \quad (\mathcal{E} - \mathcal{A})\mathbf{e}_i' = \mathbf{e}_i', \quad r+1 \leqslant i \leqslant n.$$

这意味着, $\mathcal{A}$ 和$\mathcal{E} - \mathcal{A}$ 都是投影且$\mathcal{A}^2 = \mathcal{A}$.

2.4.1 矩阵A可写成$A = (m+1)E - S$ 且

$$\det\ A = (m+1)^{n-1}(m+1-n).$$

2.4.2 分别是A_1, A_2, A_4和A_3.

2.4.3 (1)$J(\mathcal{A}) = J_2(3) \dotplus J_1(3) \dotplus J_1(3) \dotplus J_1(-2)$;

(2) 是, 当秩为1或秩为4的时候; 否, 当秩为3的时候.

2.4.4 (1)$\chi_A(t) = \chi_B(t) = (t-4)^2(t-2)$;

(2) $\mu_A(t) = (t-4)(t-2)$; $\chi_B(t) = (t-4)^2(t-2)$;

(3) $J(A) = \mathrm{diag}(2,4,4)$, $\quad J(B) = J_1(2)\dot{+}J_2(4)$.

2.4.5 研究域$\mathfrak{K}$的包含矩阵A的所有特征根的扩张域$\overline{\mathfrak{K}}$则更为方便(当然, $\mathfrak{K} = \mathbb{C} \Rightarrow \overline{\mathfrak{K}} = \mathbb{C}$). 定理4允许将A化成三角形, 又因为$\mathrm{tr}\ A^k = \mathrm{tr}(C^{-1}AC)^k$, 所以, 从一开始就可以认为A是个以$\lambda_0, \cdots, \lambda_{n-1}$ 为特征根的上三角形矩阵. 设$\lambda_0, \lambda_1, \cdots, \lambda_p$ 是所有的两两不同的特征根, 它们的重数分别是$n_0 \geqslant 0$, $n_1 \geqslant 1$, $\cdots$, $n_p \geqslant 1$ 且$\sum n_i = n$. 对于方幂A^k , 其特征根就是$\lambda_0^k, \cdots, \lambda_p^k$, 所以, 按条件

$$n_1\lambda_1^k + \cdots + n_p\lambda_p^k = 0, \quad k = 1, 2, \cdots, p.$$

当$p \geqslant 1$ 时, 我们得到一个齐次线性方程组, 它的行列式是个非零的范德蒙德行列式. 从而, $n_1 = \cdots = n_p = 0$, 矛盾. 这样一来, $\lambda_0 = 0$ 就是唯一一个n重的特征根, $\chi_A(t) = 0$, 也就是$A^n = 0$. 反命题是显然的.

2.4.6 先设A是个m阶的若尔当小块. 那么, 可以直接验证, $\Pi_m^{-1}A\Pi_m = {}^tA$. 其中$\Pi_m = \|\delta_{i,m+1-j}\|_1^m$ 是个次对角线上元素为1而其余所有元素均为零的矩阵. 如果A是若尔当块的直和, 那么, 采用Π_m 型矩阵的直和即可.

2.4.7 不失一般性, 可以认为A已经化成的若尔当标准形, $A = J(A)$. 关系式$A^N = E$ 当且仅当$J_m^N(\lambda) = E_m$. 对于$J(A)$ 上的每个若尔当块$J_m(\lambda)$ 都成立. 但这只有$m = 1$ 时才能成立. 这样一来,

$$F = J(A), A^N = E \Rightarrow A = \mathrm{diag}(\lambda_1, \cdots, \lambda_n), \lambda_i^N = 1.$$

2.4.8 先设A是个迹为零的3×3 阶幻方矩阵. 因为, $\mathrm{tr}\ A = 0 \Rightarrow \sigma(A) = 0$, 所以, 利用习题1.2.9中的矩阵$S$, 我们可以得到$AS = \sigma(A)S = 0 \Rightarrow \det\ A = 0$ (反之则$S = A^{-1}(AS) = 0$, 矛盾). 按哈密顿–凯莱定理, 我们有$A^3 = \lambda A$, 对某个$\lambda \in \mathbb{Q}$, 因为, A^3 是个幻方矩阵, 在这种情形, $A^m = \lambda^{(m-1)/2}A$ 也是个幻方矩阵, $m = 2k+1 \geqslant 3$.

现在, 如果A是任意一个3×3 阶的幻方矩阵. 那么$A_0 := A - (1/3)(\mathrm{tr}\ A)S$ 也是个幻方矩阵, 且$\mathrm{tr}\ A_0 = 0$. 由此可见, $A_0^m \in \mathrm{Mag}_3(\mathbb{Q})$, 对任意的奇数$m \geqslant 1$. 但是$SA_0 = A_0S = \sigma(A_0) = 0$, 所以

$$A^m = \left(A_0 + \frac{1}{3}(\mathrm{tr}\ A)S\right)^m = A_0^m + \left(\frac{1}{3}\mathrm{tr}\ A\right)^m S^m \in \mathrm{Mag}_3(\mathbb{Q}).$$

2.4.10 $s(t) = (t^2 - 2\mu t + \lambda\mu)/(\lambda - \mu)$, $n(t) = (t - \lambda)(t - \mu)/(\lambda - \mu)$.

2.4.11 设$\chi_{\mathcal{A}}(t) = \prod\limits_{i=1}^{p}(t-\lambda_i)^{n_i}$, 当$i \neq j$ 时, $\lambda_i \neq \lambda_j$, 且$\sum\limits_i n_i = n$. 依据定理3, 我们有根子空间的直和分解$V = \bigoplus\limits_{i=1}^{p} V(\lambda_i)$. 我们设, 对任意$\mathbf{x} \in V(\lambda_i)$, $\mathcal{S}\mathbf{x} = \lambda_i\mathbf{x}$ 且$\mathcal{N} = \mathcal{A} - \mathcal{S}$, 如果A是算子$\mathcal{A}$ 在若尔当基底之下对应的矩阵, 那么, 相应地

$$S = \lambda_1E_{n_1}\dot{+}\cdots\dot{+}\lambda_pE_{n_p}, \quad N = A - S = N_{n_1}\dot{+}\cdots\dot{+}N_{n_p},$$

其中N_{n_i} 是若尔当块$J_k(0)$, $K \leqslant n_i$ 的直和, 因为$E_{n_i}N_{n_i} = N_{n_i}E_{n_i}$, 所以$SN = NS$.

根据中国剩余定理(信以为真或到[BA Ⅲ]看看), 可得到一个复多项式$f(t)$ 使得

$$f(t)-\lambda_i=(t-\lambda_i)^{n_i}h_i(t),\quad h_i(t)\in\mathbb{C}[t],\quad i=1,\cdots,p,$$

现在, 如果$\mathbf{w}$ 是$V(\lambda_i)$ 的任意一个向量, 那么

$$\mathbf{w}=\chi_i(\mathcal{A})f_i(\mathcal{A})\mathbf{v},\quad \chi_i(t)=\prod_{j\neq i}(t-\lambda_j)^{n_j},$$

从而, $(f(\mathcal{A})-\lambda_i)\mathbf{w}=\chi_{\mathcal{A}}\mathcal{A}f_i(\mathcal{A})h_i(\mathcal{A})\mathbf{v}=\mathbf{0}$, 即, 可以令$S=f(\mathcal{A}),\ \mathcal{N}=\mathcal{A}-f(\mathcal{A})$. 显然, 可以认为$\deg f(t)<n$.

2.4.12

$$(J_n(\lambda))^k=\begin{pmatrix}\lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \cdots & \binom{k}{n-1}\lambda^{k-n+1}\\ 0 & \lambda^k & k\lambda^{k-1} & \cdots & \binom{k}{n-2}\lambda^{k-n+2}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda^k\end{pmatrix}$$

2.4.13 根据哈密顿–凯莱定理(同样可见[BA Ⅰ]的第二章末尾的习题14), 2×2 阶矩阵$[X,Y]=XY-YX$ 满足关系式$[X,Y]^2=\lambda E,\ \lambda=-\det[X,Y]$. 从而, 正如我们已知的, $\operatorname{tr}[X,Y]=0$. 因为$[E,Z]=0$, 所以, $\left[[X,Y]^2,Z\right]=0$.

3.1.1 (1) $\langle t^2-1/3\rangle$.

3.1.2 (1) $\cos\alpha=3/\sqrt{10}$;

(2) $\mathbf{z}\perp\mathbf{x}\Rightarrow z_1+z_3=0,\quad \mathbf{z}=[z_1,\ z_2,\ z_3]$.

3.1.3 设$(\mathbf{e}_1,\cdots,\mathbf{e}_n)$ 是n维向量空间V 的一个标准正交基底. 若向量组$(\mathbf{a}_1,\cdots,\mathbf{a}_n)$ 可表成$\mathbf{a}_k=\sum\limits_{i=1}^{n}\mathbf{a}_{ik}\mathbf{e}_i$, 那么, 当矩阵$A=(a_{ij})$ 非退化时, 这个向量组也是V 的一个基底. 遵照定理6, 将它们施以正交化过程, 我们就得到一个标准正交基底$(\mathbf{e}'_1,\cdots,\mathbf{e}'_n)$ 其中$\mathbf{e}'_j=\sum\limits_{k=1}^{j}c'_{kj}\ \mathbf{a}_k,\quad c'_{kj}\neq 0$; 转换矩阵$C'=(c'_{kj})$ 是上三角形矩阵. 它的逆矩阵$C=(C')^{-1}=(c_{kj})$ 同样也是个上三角形矩阵. 我们有

$$\mathbf{e}'_j=\sum_k c'_{kj}\mathbf{a}_k=\sum_i\left(\sum_k a_{ik}c'_{kj}\right)\mathbf{e}_i=\sum_i b_{ij}\mathbf{e}_i.$$

作为一个从一个标准正交基底$\mathbf{e}_i$ 向另一个标准正交基底$\mathbf{e}'_i$ 的转换矩阵, 矩阵$B=(b_{ij})$ 是正交矩阵. 这样一来, $B=AC'=AC^{-1}$, $A=BC$.

3.1.4 矩阵$(E\pm A)^{\pm 1}$, $A^{\pm}$ 是可交换的. 所以

$$\begin{aligned}{}^tK&={}^t(E-A)^{-1}\cdot{}^t(E+A)=(E-{}^tA)(E+{}^tA)\\&=(E-A^{-1})(E+A^{-1})=(E-A^{-1})A^{-1}\cdot A(E+A^{-1})\\&=(A(E-A^{-1}))^{-1}(E+A)=(A-E)^{-1}(E+A)=-K.\end{aligned}$$

$${}^tA = {}^t(E-K)^{-1t}(E+K) = (E-{}^tK)^{-1}(E+{}^tK) = (E+K)^{-1}(E-K) = A^{-1}.$$

还要注意到, $\det(E-A)=0 \Leftrightarrow$ 对任意$\mathbf{0}\neq\mathbf{x}\in M_n(\mathbb{R})$ 都有$(E-A)\mathbf{x}=\mathbf{0}$.

3.1.5 由(18)推出

$$\begin{aligned}\det\ A &= \det(E-K)^{-1}\det(E+K) = (\det(E-K))^{-1}\det{}^t(E+K)\\ &= (\det(E-K))^{-1}\det(E+{}^tK) = (\det(E-K)^{-1})\det(E-K) = 1.\end{aligned}$$

类似地, 我们进入(19)情形.

3.1.6 由正交矩阵A的定义得到, $\phi_A(\lambda)=0$, 所以, 又有$\phi_A(1/\lambda)=0$, 这是因为$1/\lambda=\overline{\lambda}/(\lambda\overline{\lambda})=\overline{\lambda}$, 而$\phi_A(t)$ 是实多项式. 可见, 多项式$f(t):=t^n\phi_A(1/t)$ 和$\phi_A(t)$ 具有相同的根, 且重数也一样. 剩下来只需补充, $\phi_A(0)=\pm1$,所以多项式$f(t)$ 和$\phi_A(t)$ 的最高系数只差一个符号.

3.1.7 化成空间$\mathbb{R}^n$ 的标准正交基底

$$(A_{(1)}/||A_{(1)}||,\cdots,A_{(n)}/||A_{(n)}||)$$

3.1.8 由$(X_{(i)})$ 得到空间$\mathbb{R}^n$ 的一个正交基底$(Y_{(i)})$, 比较$||Y_{(i)}||$和$||X_{(i)}||$, 行列式$\det[Y_{(1)},\cdots,Y_{(n)}]$ 和$\det[X_{(1)},\cdots,X_{(n)}]$, 再利用前面的练习题.

3.2.1 不是, 只要取矩阵

$$A=\begin{pmatrix} i & -\sqrt{2}\\ \sqrt{2} & i\end{pmatrix}$$

为例, 我们有${}^tA\cdot A=E$, $\det\ A=1$, 也就是$A\in S\mathbb{C}O(n)$, 但${}^t\bar{A}\cdot A\neq E$, 所以, $A\notin U(n)$.

3.3.1 不失一般性, 可以认为酉矩阵$A\in SU(n)$ 有对角形式

$$A=\mathrm{diag}(\lambda_0,\lambda_1,\cdots,\lambda_{n-1}),\quad \lambda_0=\overline{\lambda_1}\cdots\overline{\lambda_{n-1}},\quad \lambda_i\overline{\lambda_i}=1.$$

当$n>2$ 时, 酉矩阵对子

$$X=\begin{pmatrix}0&0&\cdots&0&1\\1&0&\cdots&0&0\\0&1&\cdots&0&0\\ \cdots&\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1&0\end{pmatrix},\quad Y=\begin{pmatrix}0&\alpha_1&0&\cdots&0\\0&0&\alpha_2&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&\alpha_{n-1}\\1&0&0&\cdots&0\end{pmatrix},\quad \alpha_i\overline{\alpha_i}=1.$$

是方程式$XYX^{-1}Y^{-1}=A$ 的一个解.

可以直接验证, $\alpha_k=\lambda_k\lambda_{k+1}\cdots\lambda_{n-1}$ 是个完全渐近值. 当$n=2$ 时, 我们有

$$A=\begin{pmatrix}\lambda&0\\0&\overline{\lambda}\end{pmatrix},\quad X=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad Y=\begin{pmatrix}0&\alpha\\-\overline{\alpha}&0\end{pmatrix},$$

其中$\lambda = e^{i\theta}$, $\alpha = e^{i(\pi-\theta)/2}$.

3.3.2 设$\chi_n(t) = \chi_J(t)$ 是n阶雅可比矩阵的特征多项式($\chi_0(t)$)．考察J的主子式并按最后一行的元素分解, 我们就得到一个递归关系式

$$\chi_k(t) = (a_k - t)\chi_{k-1}(t) - b_{k-1}c_{k-1}\chi_{k-2}(t), \quad k \geqslant 2.$$

由此可得, $\chi_n(t)$ 由乘积$b_1c_1, \cdots, b_{n-1}c_{n-1}$ 决定而不是由数b_k, c_k 单独地决定. 在这种情况下, 用$\sqrt{b_kc_k}$ 代替J的系数b_k，c_k，我们得对称的雅可比矩阵J'，它具有同样的特征多项式$\chi_n(t)$．但是, 正如我们已经知道的(见附录§2), 矩阵J' 的特征根都是实的, 所以, $\mathrm{Spec}(J') \subset \mathbb{R}$．

转向任意特征值λ 的重数问题(J作用在$\mathbb{R}^n$ 上), 注意, 特征向量$\mathbf{x}$ 的坐标$x_1, \cdots, x_n$ 由线性关系式

$$-c_{k-2}x_{k-2} + (a_{k-1} - \lambda)x_{k-1} - b_{k-1}x_k = 0,$$

$$k = 2, 3, \cdots, n+1 \ (c_0 = 0 = b_n).$$

联系起来, 按照这些关系式, $x_k = \sigma_k x_1$，其中比例系数σ_k 具有分数形式: 分母是$b_1b_2\cdots\ b_{k-1}$，分子是$b_1c_1, \cdots, b_{k-1}c_{k-1}, (a_1 - \lambda), \cdots, (a_{k-1} - \lambda)$ 的函数. 不失一般性, 可以认为$x_1 = 1$．此时, $\mathbf{x}$ 就是唯一确定的特征向量. 可见, 根λ 的代数重数, 在这种情形, 和它的几何重数一致, 等于1.

3.3.3 是.

3.3.4 显然, 同时对角化的算子$\mathcal{A}$, $\mathcal{B}$ 是可交换的. 反过来, 设$\mathcal{A}$, $\mathcal{B}$ 是可交换的对角的线性算子. $\mathcal{A}$ 的对角性意味着$V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_p}$，$\{\lambda_1, \cdots, \lambda_p\} = \mathrm{Spec}(\mathcal{A})$. 根据引理6, 其中$\mathbb{C}$ 可以用任意域来代替, 只要该域包含$\mathrm{Spec}(\mathcal{A})$和$\mathrm{Spec}(\mathcal{B})$, 而每个子空间$V^{\lambda_i}$ 对$\mathcal{B}$ 都是不变的. 但是, 如果W 是任意一个不变子空间, 那么$\mathcal{B}$ 的对角性质导致它在W 的限制的对角性. 特别地, 在V^{λ_i} 可选取基底$(\mathbf{e}_1^{(i)}, \cdots, \mathbf{e}_{n_i}^{(i)})$，在此基之下, $\mathcal{B}$ 取对角形式. 因为对$i = 1, \cdots, p$ 都成立, 又因为$\mathcal{A}\mathbf{e}_k^{(i)} = \lambda_i\mathbf{e}_k^{(i)}$，所以, $(\mathbf{e}_1^{(1)}, \cdots, \mathbf{e}_{n_1}^{(1)}, \mathbf{e}_1^{(2)}, \cdots, \mathbf{e}_{n_p}^{(p)})$ 是个基底, $\mathcal{A}$ 和$\mathcal{B}$ 在此基底之下可对角化.

3.3.5 因为$\mathcal{AB} = \mathcal{BA}$, 由$\mathcal{A}$ 和$\mathcal{B}$ 的埃尔米特性(对称性)可推出乘积的埃尔米特性. 除此以外, 由谱分解定理推出, $\sqrt{\mathcal{A}}$ 是$\mathcal{A}$ 的多项式, $\sqrt{\mathcal{B}}$ 是$\mathcal{B}$ 的多项式. 因为

$$\mathcal{AB} = \mathcal{BA} \Rightarrow \sqrt{\mathcal{A}}\sqrt{\mathcal{B}} = \sqrt{\mathcal{B}}\sqrt{\mathcal{A}},$$

所以, 我们有

$$\mathcal{AB} = (\sqrt{\mathcal{A}}\sqrt{\mathcal{A}})(\sqrt{\mathcal{B}}\sqrt{\mathcal{B}}) = (\sqrt{\mathcal{A}}\sqrt{\mathcal{B}})^2,$$

又因为$(\sqrt{\mathcal{A}})^* = \sqrt{\mathcal{A}} > 0$，$(\sqrt{\mathcal{B}})^* = \sqrt{\mathcal{B}} > 0$，那么, $(\sqrt{\mathcal{A}}\sqrt{\mathcal{B}})^* = \sqrt{\mathcal{A}}\sqrt{\mathcal{B}}$，进而, $\mathcal{AB} > 0$.

3.3.6 显然, ${}^t(A^2) = ({}^tA)^2 = (-A)^2 = A^2$，所以, 可以把$A$解释成带有标准纯量积的列向量空间上的线性算子, 我们有$(A^2\mathbf{x}|\mathbf{x}) = (A\mathbf{x}|{}^tA\mathbf{x}) = -(A\mathbf{x}|A\mathbf{x}) \leqslant 0$. 其次, $A\mathbf{x} = \lambda\mathbf{x} \Rightarrow A^2\mathbf{x} = \lambda^2\mathbf{x}$ 且$\lambda^2 \leqslant 0$．

3.3.7 考察两个二次型

$$q(\mathbf{x}) = (\mathcal{A}\mathbf{x}|\mathbf{x}), r(\mathbf{x}) = (\mathcal{B}\mathbf{x}|\mathbf{x}),$$

它们分别由$\mathcal{A}$ 和$\mathcal{B}$ 对应, 按条件$q(\mathbf{x})$ 是正定的, 所以, 按§2的定理8, 二次型$q(\mathbf{x})$, $r(\mathbf{x})$ 可以同时化成规范型. 如果算子$\mathcal{A}$, $\mathcal{B}$ 在相应的基底之下的矩阵是$A = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$, $B = \operatorname{diag}(\mu_1, \cdots, \mu_n)$, 那么

$$\operatorname{Spec}(\mathcal{A}\mathcal{B}) = \operatorname{Spec}(AB) = \{\lambda_1\mu_1, \cdots, \lambda_n\mu_n\} \in \mathbb{R},$$

因为$\lambda_i \in \mathbb{R}$, $\mu_i \in \mathbb{R}$. 也可以用其他办法证明.

3.3.8 为了决定二次型$q(\mathbf{x})$ 的平稳值, 利用分析学中拉格朗日方法是很方便的. 在欧几里得向量空间V 的直角坐标系之下应有$q(\mathbf{x}) = \sum\limits_{i,j} f_{ij}x_ix_j$, $(\mathbf{x}|\mathbf{x})$ =1. 与拉格朗日方法相对应, 建立函数

$$L(x_1, x_2, \cdots, x_n) = \sum_{i,j=1}^{n} f_{ij}x_ix_j - \lambda\sum_{i=1}^{n} x_i^2$$

且按x_i , $i = 1, \cdots, n$ 求一阶偏微商并取零, 我们得到一个线性方程组

$$\sum_{j=1}^{n} (f_{ij} - \delta_{ij}\lambda)x_j = 0, \quad i = 1, 2, \cdots, n,$$

这个线性方程组在我们求对称算子$\mathcal{F}$ 的特征值和特征向量的时候已经遇到过, $\mathcal{F}$ 就是与q相伴的那个算子. 这就得到了所需的断言.

换言之, 按证明, 二次型$q(\mathbf{x})$ 在算子$\mathcal{F}$ 的某个特征向量$\mathbf{e}_i$ (长度为1)处取得稳定值. 因为$\mathcal{F}\mathbf{e}_i = \lambda_i\mathbf{e}_i$, 所以

$$q(\mathbf{e}_i) = (\mathbf{e}_i|\mathcal{F}\mathbf{e}_i) = \lambda_i(\mathbf{e}_i|\mathbf{e}_i) = \lambda_i.$$

这意味着, 二次型$q(\mathbf{x})$ 的稳定值和它的规范型的系数是一致的.

3.3.9 提示. 不失一般性, 可以把$\mathbb{C}$ 上的向量空间V , $\dim\ V < \infty$ 上的可交换的算子族看成是有限的, 原因是, 由于$\mathcal{L}(V)$ 的维数有限, 这个算子族中, 总可以选出有限基的子集

$$\{\mathcal{A}_1, \cdots, \mathcal{A}_m | \mathcal{A}_i\mathcal{A}_j = \mathcal{A}_j\mathcal{A}_i, 1 \leqslant i, j \leqslant m\}.$$

然后, 对m归纳, 再用引理6的证明中推断.

如果, 已经找到一个特征向量$\mathbf{x}$, 对$\mathcal{A}_1, \cdots, \mathcal{A}_{m-1}$ 有$\mathcal{A}_i\mathbf{x} = \lambda_i\mathbf{x}$, 那么, 看$\mathcal{A}_m$ 不变的子空间$W = \mathbb{C}[\mathcal{A}_m]\mathbf{x}$, 而且, 在这个子空间中, 特征向量$\mathbf{y} = f(\mathcal{A}_m)\mathbf{x} : \mathcal{A}_m\mathbf{y} = \lambda_m\mathbf{y}$, 其中$f$是某个多项式. 但是, 在这种情况下, $\mathcal{A}_i\mathbf{y} = \mathcal{A}_i f(\mathcal{A}_m)\mathbf{x} = f(\mathcal{A}_m)\mathcal{A}_i\mathbf{x} = f(\mathcal{A}_m)\lambda_i\mathbf{x} = \lambda_i\mathbf{y}$. $1 \leqslant i \leqslant m-1$.

3.3.10 提示. 当n=1时, 命题显然.然后对n用归纳法.可以认为, $A_j, j \in J$ 是作用在向量空间V上的线性算子$\mathcal{A}_j$ 在固定的基底$(\mathbf{e}_1, \cdots, \mathbf{e}_n)$之下的矩阵, 同时, $\mathcal{A}_i\mathcal{A}_j = \mathcal{A}_j\mathcal{A}_i$, $i, j \in J$. 依据习题3.3.9, 算子族$\{\mathcal{A}_j | j \in J\}$ 有公共的特征向量$\mathcal{A}_j\mathbf{x} = \lambda_j\mathbf{x}$. 不失一般性, 可以认为$\mathbf{x} = \mathbf{e}_1$, 如有必要, 可用共轭集合$\mathfrak{G}' = (C')^{-1}\mathfrak{G}C'$ 来代替$\mathfrak{G}$. 在这样约定之下, 我们就有

$$A_j = \begin{pmatrix} \lambda_j & \alpha_{j2} \cdots \alpha_{jn} \\ 0 & B_j \end{pmatrix}, \quad B_j \in M_{n-1}(\mathbb{C}). \tag{$*$}$$

由条件$A_iA_j = A_jA_i$ 推出$B_iB_j = B_jB_i$, 而且按归纳假设, 必可找到一个非退化矩阵$D \in M_{n-1}(\mathbb{C})$ 使得所有的矩阵$D^{-1}B_iD$ 都是上三角形的.现在, 只要令

$$C = C'C'', \quad C'' = \begin{pmatrix} 1 & 0 \\ 0 & D \end{pmatrix}$$

就足够了.

3.3.12 提示. 与习题3.3.11相对应, $M_n(\mathbb{C})$ 的可交换的子代数的最大维数不能小于$[n^2/4]+1$. 为了证明这个数目的极大性, 我们利用文章: Mirzakhani M.//Amer. Math.Monthly.—1998.—March—P.260—262的一个结论. 我们设$m \geqslant [n^2/4] + 2$. 如同在习题3.3.10一样, 把A_j , $1 \leqslant j \leqslant m$ 记成$(*)$形式且$B_j \in M_{n-1}(\mathbb{C})$ 是可交换的: $B_iB_j = B_jB_i$.设

$$\mathfrak{K} = \langle B_1, \cdots, B_m \rangle_{\mathbb{C}}.$$

按归纳法假定$r = \dim \mathfrak{K} \leqslant [(n-1)^2/4] + 1$. 不失一般性, 我们可以认为$B_1, \cdots, B_r$ 是线性无关的, 从而$B_i = \sum_{j=1}^{r} \alpha_{ij}B_j$. 当$i > r$ 时, 我们设$C_i = A_i - \sum_{j=1}^{r} \alpha_{ij}A_j$. 显然, 矩阵组$C_j$ 是线性无关的, 且当$i = r+1, \cdots, m$ 时, 它们中的每一个都形如$C_i = \begin{pmatrix} \mathbf{c}_i \\ 0 \end{pmatrix}$, 其中$\mathbf{c}_i$ 是$1 \times n$阶矩阵.此外, 向量$\mathbf{c}_i$ 在$\mathbb{C}$ 上应该是线性无关的.

现在, 我们注意, 上三角形矩阵A_j 同样可以记成

$$A_j = \begin{pmatrix} B'_j & \mu_{1j} \\ & \vdots \\ 0 & \mu_{nj} \end{pmatrix}$$

的形状, B'_j 是交换矩阵.再次利用归纳假设, 我们就导出一个线性无关的$n \times 1$阶矩阵的$\mathbf{c}'_{s+1}, \cdots, \mathbf{c}'_m$ 的集合, 其中$s \leqslant [(n-1)^2/4] + 1$. 此时, $C'_j = \begin{pmatrix} \mathbf{c}'_j \\ 0 \end{pmatrix}, j \geqslant s+1$.
由于C_i 和C'_j 属于同一个可交换族, 所以, $\mathbf{c}_i\mathbf{c}'_j = 0$对任意$i = r+1, \cdots, m$ 和$j = s+1, \cdots, m$ 都成立.

最后, 我们考察$(m-r)\times n$ 阶矩阵C, 它的第i行上是$\mathbf{c}_i, i=r+1,\cdots,m$.因为$\mathbf{c}_i$线性无关, 所以rank $A\geqslant m-r$.另一方面, $C\mathbf{c}'_j=0$ 对所有$j=s+1,\cdots,m$ 都成立. 因为$\mathbf{c}'_j$ 线性无关又因为对矩阵C的线性算子$\mathcal{C}$都有dim Ker $\mathcal{C}$ + dim Im$\mathcal{C}=n$(第二章§1的定理4), 所以, 借助对于m,r 和s的不等式, 我们就有

$$n\geqslant(m-r)+(m-s)\geqslant 2\left(\left[\frac{n^2}{4}\right]-\left[\frac{(n-1)^2}{4}\right]+1\right)\geqslant 2\left[\frac{n}{2}\right]+2>n,$$

这个矛盾完成了我们所需要的证明.

3.3.13 提示. 在V 中选一个基底, 在此基底之下二次型f对应的矩阵是

$$J_0=\begin{pmatrix}0 & -E_m\\ E_m & 0\end{pmatrix}.$$

研究以

$$\Phi=\begin{pmatrix}0 & -E_m\\ -E_m & 0\end{pmatrix}$$

为矩阵的非退化的对称型$\varphi(\mathbf{x},\mathbf{y})$, 并按照这个矩阵做分解$V=V_1\oplus V_2$, 因此,

$$f(\mathbf{x},\mathbf{y})=\varphi(\mathbf{x}_1,\mathbf{y}_2)-\varphi(\mathbf{x}_2,\mathbf{y}_1).$$

另一方面, 因为在线性算子和双线性的对应中有

$$\varphi(\mathbf{x},\mathbf{y})=(\mathbf{x}|\mathcal{A}\mathbf{y}),$$

其中$\mathcal{A}$ 是线性算子, 它具有所需的所有的性质 : φ 的非退化性$\Rightarrow$ $\mathcal{A}$ 的非退化性; φ 的对称性$\Rightarrow$ $\mathcal{A}$ 的对称性.

3.4.1 考察一个标准正交基底, 在这个基底之下$\mathcal{A}$ 取规范形式, 且把每个小块

$$\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

看成是空间U的一维子空间上用$e^{i\varphi}$ 相乘.

3.4.2 设$\lambda_j=\alpha_j+i\beta_j$ (复数的通常记法). 那么, 在基底$(\mathbf{e}_1,i\mathbf{e}_1,\cdots,\mathbf{e}_n,i\mathbf{e}_n)$之下算子$\mathcal{A}_{\mathbb{R}}$ 的矩阵$\mathrm{A}_{\mathbb{R}}$ 沿对角线取

$$\begin{pmatrix}\alpha_j & -\beta_j\\ \beta_j & \alpha_j\end{pmatrix}$$

的分块三角形的形式. 按照角上为零的行列式的计算规则, 应有

$$\det A_{\mathbb{R}}=\prod_{j=1}^{n}(\alpha_j^2+\beta_j^2)=\prod_{j=1}^{n}|\lambda_j|^2=\left|\prod\lambda_j\right|^2=|\det A|^2.$$

由于线性算子的行列式与基底的选择无关, 所以, 这就证明了断言. 剩下来要再实现一些细节了.

3.5.1 因为级数$\sum_{k=1}^{\infty}(1/\sqrt{k})^2$ 是发散的.

3.5.7 对任意$T_n(t)$ 有$t = \cos(2i-1)\dfrac{\pi}{n+1}$; 且对任意$U_n(t)$ 有$t = \cos\dfrac{i\pi}{n+1}$; $i = 1, 2, \cdots, n$.

3.5.8 是的, 准确到变量的线性替换, f_n 是第二类切比雪夫多项式.

4.1.1 按照定理4的推论

$$\Pi' \cap \Pi'' = \dot{r} + U, \quad U = U' \cap U''.$$

因此

$$\dot{r} \in \Pi' \cap \Pi'' \Leftrightarrow \dot{p} + \mathbf{u}' = \dot{r} = \dot{q} + \mathbf{u}'' \Leftrightarrow \overrightarrow{pq} = \mathbf{u}' - \mathbf{u}'' \in U' + U''.$$

4.1.6 只要注意到断言对于等边三角形成立就足够了, 因为存在仿射变换f, 它把等边三角形变成给定的三角形, 且f把直角变成直角, 中点变成中点, 从而把中线变成中线.

4.2.1 7.

4.2.2 7.

4.3.1 对称群S_3 .

4.3.2 绕点$(-1/\sqrt{2},\ 1 + 1/\sqrt{2}\ )$旋转$\pi/4$ 角.

4.3.3 (1) 两个正交平面上绕不动点的旋转.

(2) 一个平面旋转加一个正交平面平移;

(3) 平移.

5.2.3 如同在习题3.3.8一样, 相应于拉格朗日方法, 建立函数$\sum_i x_i^2 - \lambda\sum_{i,j} f_{ij}x_ix_j$, 且给出极值性条件$x_i - \lambda\sum_j f_{ij}x_j = 0$, $1 \leqslant j \leqslant n$. 化成对于$\mu = 1/\lambda$ 的特征方程.

5.2.4 $-1/2 < t < 1$.

5.2.5 双曲线.

5.2.6 在它们的方程式的所有的对应的系数都成比例的情形(这是可能的, 除非是自由的).

6.1.3 提示. 见[2]第4章§1.

6.1.4 例如, 取$V = \mathfrak{K}^2 = W$ 并考察元素.

$$(1, 0) \otimes (0, 1) + (0, 1) \otimes (1, 0).$$

如果试图将它表达成$(a, b) \otimes (c, d)$ 的形式, 那么, 就会最终得到一些矛盾的关系式$ad = 1, bc = 1, ac = 0, bd = 0$.

6.1.5 $(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$.

6.1.6 两个矩阵与$A\otimes B$ 重合.

6.3.2 提示. 利用关系式(10).

6.3.3 提示. 利用拉普拉斯公式[BA I]第3章§3习题8.

7.1.4 由习题7.1.3得到一个对于代数$\mathfrak{su}(2)$的有用的自同构映射. 实际上, 如果

$$A=\alpha_1T_1+\alpha_2T_2+\alpha_3T_3,\quad B=\beta_1T_1+\beta_2T_2+\beta_3T_3,$$

那么,

$$[A,B]=AB-BA=C=\gamma_1T_1+\gamma_2T_2+\gamma_3T_3.$$

进而我们容易验证, 向量

$$(\gamma_1,\gamma_2,\gamma_3)=(\alpha_1,\alpha_2,\alpha_3)\times(\beta_1,\beta_2,\beta_3)$$

是个向量积(或外积).

同样, 对$\mathfrak{so}(3,\mathbb{R})$可以确信, 如果注意到, 矩阵

$$P_1=\begin{pmatrix}0&0&1\\0&0&-1\\0&1&0\end{pmatrix},P_2=\begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix},P_3=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}$$

构成空间$\mathfrak{so}(3,\mathbb{R})$的一个基底, 且

$$[P_1,P_2]=P_3,\quad [P_3,P_1]=P_2,\quad [P_2,P_3]=P_1.$$

7.1.5 由显然的关系式

$$(B^{-1}AB)^k=B^{-1}AB\cdot B^{-1}AB\cdots B^{-1}AB=B^{-1}A^kB$$

可推出

$$\begin{aligned}\exp\,(B^{-1}AB)&=\sum_{k\geqslant 0}\frac{1}{k!}(B^{-1}AB)^k=\sum_{k\geqslant 0}\frac{1}{k!}B^{-1}A^kB\\&=B^{-1}\cdot\sum_{k\geqslant 0}\frac{1}{k!}A^k\cdot B=B^{-1}(\exp A)\cdot B.\end{aligned}$$

7.1.6 $(e^2/2)A$.

7.1.7 由第3章§3已经知道, 在标准正交基底之下, 酉算子$\mathcal{A}$ 的矩阵A 可以记成$A=B\cdot\mathrm{diag}\{e^{i\varphi_1},\cdots,e^{i\varphi_n}\}\cdot B^{-1}$, 所以, $A=\exp\,i(B\mathrm{diag}\{\varphi_1,\cdots,\varphi_n\}B^{-1})$.

7.1.8 如同对所有的保距映射一样, $||\mathcal{A}\mathbf{x}||=||\mathbf{x}||$, 所以, $||(\mathcal{AB})\mathbf{x}||=||\mathcal{A}(\mathcal{B}\mathbf{x})||=||\mathcal{B}\mathbf{x}||$, 剩下来就是显然的了:

$$||\mathcal{AB}||=\sup_{||\mathbf{x}||=1}||(\mathcal{AB})\mathbf{x}||=\sup_{||\mathbf{x}||=1}||\mathcal{B}\mathbf{x}||=||\mathcal{B}||.$$

7.1.9 选择一个基底, 在这个基底之下, 正规算子$\mathcal{A}$ 有对角矩阵

$$A=\operatorname{diag}\{\lambda_1,\cdots,\lambda_n\},\qquad |\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|.$$

所以

$$||A||=\sup_{|x_1|^2+\cdots+|x_n|^2=1}\sqrt{|\lambda_1x_1|^2+\cdots+|\lambda_nx_n|^2}=|\lambda_1|.$$

从而可推出整个断言.

7.1.10 推断纲要在[15]的叙述中已经足够自然了. 因为$r(\mathcal{A})^k\leqslant||\mathcal{A}^k||$ (见第9小节的2)), 所以, 在$\lim\limits_{k\to\infty}\mathcal{A}^k=\mathcal{O}$ 的情形, 我们有$\lim\limits_{k\to\infty}r(\mathcal{A})^k=0$. 而这就意味着, $r(\mathcal{A})<1$. 反过来, 当$r(\mathcal{A})=1-2\varepsilon<1$ 且k充分大的时候, 由第6小节的1)推出, $||\mathcal{A}^k||^{1/k}\leqslant r(\mathcal{A})+\varepsilon=1-\varepsilon$, 从而$||\mathcal{A}||^k\leqslant(1-\varepsilon)^k$ 且由此可见, $\lim\limits_{k\to\infty}\mathcal{A}^k=\mathcal{O}$.

7.3.1 设

$$S=\{\dot a|f_i(\dot a)\geqslant0,i=1,\cdots,m\}$$

是我们的有界多边形. 如果$\dot a\in S$, 又比如说,

$$f_1(\dot a)=0,\cdots,f_r(\dot a)=0,\ f_{r+1}(\dot a)>0,\cdots,f_m(\dot a)>0,\tag{$*$}$$

那么, 方程组$f_i(\dot x)=0,1\leqslant i\leqslant r$, 决定了平面$\prod\limits_{\dot a}$ (当r=0时, $\prod\limits_{\dot a}=S$). 集合$S_{\dot a}=\prod\limits_{\dot a}\cap S$ 就是S 中包含点$\dot a$ 的边界.

现在, 设, S' 是多边形S 的所有顶点的凸包络. 因为S 是个凸集, 所以, $S'\subset S$. 剩下来就是要说明所有的点$\dot a\in S$ 必然包含在S' 中. 我们对$\dim\prod\limits_{\dot a}$ 用归纳法来证明这件事. 如果$\dim\prod\limits_{\dot a}=0$, 那么, $\dot a$ 是个顶点, 且由此可见, 按定义, 它包含在S' 中. 进而, 我们可以认为, $\dim\prod\limits_{\dot a}>0$, 为确定起见, 可设关系式$(*)$被满足. 我们在平面$\prod\limits_{\dot a}$ 上经过点$\dot a$ 引出任意直线$\{\dot a+\lambda\otimes\}$, 它与多边形S 的交集, 由不等式组

$$f_i(\dot a+\lambda\mathbf{x})=f_i(\dot a)+\lambda\mathcal{F}_i(\mathbf{x})\geqslant0,\quad r+1\leqslant i\leqslant m$$

给出($\mathcal{F}_i$是函数f_i 的线性部分). 多边形S 有限, 所以, 给出的交集可表成一个线段$\dot p\dot q$. 就如同在全体平面$\prod_{\dot a}$ 上一样, 函数$f_1,\cdots,f_r$ 在点$\dot p$, $\dot q$ 的每一处都转化为零. 但是, 即使是对某一个$i>r$ 有$f_i(\dot p)=f_i(\dot q)=0$, 就意味着$\dim\prod_{\dot p}<\dim\prod_{\dot a}$ 且$\dim\prod_{\dot q}<\dim\prod_{\dot a}$. 按归纳假设, 我们可以认为, $\dot p$, $\dot q\in S'$. 因为$\dot a\in\dot p\dot q$, 所以, 就有$\dot a\in S'$. 这个论断, 显然, 差不多重复了第7章§3的定理的证明.

教 法 说 明

质朴地想一下, 本书的所有内容都曾经在真正受读者欢迎的教科书中叙述过了. 实际上, 从第3章开始就已经不得不放弃个别的片断, 而第7章只能部分地触及(线性算子的指数函数和罗巴切夫斯基几何学). 确实, 某些材料被有意地挪到习题里了. 下面的口试测试问题清单反映出教科书的一个版本. 每个考签包含两个取自不同章节的问题, 而且并不十分难于解答.

测试的问题

1. 关于域上有限维向量空间基底的定理.
2. 在向新的基底转化时向量坐标的替换规律.
3. 具有同一个有限维数的空间的同构.
4. 关于子空间之和的维数的定理.
5. 何时子空间之和是个直和?
6. 关于对偶向量空间维数的定理. 自反性.
7. 齐次线性方程组解的几何解释.
8. 用矩阵给定向量空间的线性映射. 向量的坐标变换.
9. 用术语“核”刻画线性映射的双射性(再用“像”的术语刻画).
10. 线性算子的代数. 极小多项式.刻画算子的非退化性.
11. 关于在不同基底之下算子的矩阵之间关系的定理.
12. 不变子空间：一般事实；关于投影算子的定理.
13. 特征向量和特征值. 特征多项式.
14. 关于几何重数和代数重数的定理. 算子迹的性质.
15. 关于有素谱的线性算子可对角化的定理.

16. 复线性算子和实线性算子的不变子空间.
17. 关于把复线性算子化成三角形的定理.
18. 哈密顿–凯莱定理及其推论.
19. 叙述关于矩阵若尔当标准型的定理及其推论(可对角化的判别方法).
20. 关于幂零矩阵的若尔当标准型.
21. 关于把空间分解为根子空间的直和的定理.
22. 矩阵若尔当标准型的唯一性.
23. 在不同基底之下的双线性型的矩阵.
24. 对称的与斜对称的双线性型. 二次型.
25. 关于把对称双线性型化成规范型的定理.
26. 实二次型的符号差的唯一确定性(惯性定律).
27. 化非退化的对称双线性型的雅可比方法.
28. 正定型和正定矩阵. 西尔维斯特判别法.
29. 斜对称的双线性型的规范型.
30. 欧几里得向量空间. 柯西–布尼亚科夫斯基不等式及其推论.
31. 标准正交基底的存在定理. 格拉姆–施密特过程.
32. 关于空间正交分解的定理.
33. 欧几里得向量空间和对偶空间的自然同构.
34. 标准正交基底和正交矩阵. 群$O(n)$ 和群$SO(n)$.
35. 欧几里得向量空间上线性算子与双线性型的关系. 自共轭性质.
36. 关于自共轭算子的可对角化定理.
37. 化二次型到主轴上去. 矩阵的表述.
38. 关于将两个二次型同时演化的定理.
39. 关于正交算子的规范型矩阵的定理.
40. 关于把非退化算子表示成自共轭算子和正交算子的复合形式的定理.
41. 埃尔米特型和埃尔米特空间.标准正交基底的存在性.
42. 埃尔米特算子和酉线性算子. 群$U(n)$ 和群$SU(n)$.
43. 仿射空间；同构；坐标系.
44. 仿射线性函数和线性方程组. 子空间的表达.
45. 关于欧几里得空间中点到平面的距离定理.
46. 最小二乘法. 关于函数逼近的概念.超定的线性系统.
47. 关于欧几里得空间中平面间的距离的定理.
48. 格拉姆行列式和平行六面体的体积. 实空间中关于方向的概念.
49. 把欧几里得点空间上的仿射变换分解成平移, 有不动点的运动以及沿相互垂直的方向伸缩的乘积.
50. 作为有向体变换系数的仿射变换的行列式.

51. 直线上和平面上运动的分类.
52. 3维欧几里得空间的运动的分类.
53. 仿射空间上的二次函数. 中心的性质.
54. 把二次函数化成规范型.
55. 二次曲面与二次函数之间的对应.
56. 二次曲面的类型.
57. 二次曲面的大秩和小秩. 锥面与柱面.
58. 张量概念. 价$\leqslant 2$的张量.
59. 张量的坐标. 关于卷积的概念.
60. 斜对称张量. 交错算子的性质.
61. 外积和外代数.
62. 外代数的基底.
63. 外积与行列式.

索　引

A

B

C

D

E

F

G

Y

Z

其他